DOCUMENT RESUME

ED 071 196	95	EA 004 804

TITLE Incentives in Education Project, Impact Evaluation Report. Final Report.
INSTITUTION Planar Corp., Washington, D.C.
PUB DATE 24 Oct 72
CONTRACT OEC-0-71-4770
NOTE 398p.

EDRS PRICE MF-$0.65 HC-$13.16
DESCRIPTORS *Achievement Gains; Demonstration Projects; Disadvantaged Groups; Evaluation Needs; *Incentive Systems; Models; Parent Attitudes; *Parent Participation; Parent School Relationship; *Performance Criteria; Policy Formation; Statistical Analysis; Student Attitudes; Teacher Attitudes; *Teacher Motivation

ABSTRACT

This report describes results of a demonstration project carried out in four cities during 1971-72. The project aimed at exploring the feasibility and impact of two different forms of money incentives payments. In one form -- the "Teacher-Only" model -- the teachers in a school were offered a series of bonuses ranging from $150 to $600 per class per subject, depending on the amount of gain shown by their students on standardized tests of reading and math. In the "Parent-Teacher" model, the same bonus offer was made to teachers; but in addition, there were cash payments of $12.50 to $50 offered to each parent, depending on the mean gain shown by all of the students in their child's class. Quantitative data were collected from students, teachers, and parents at two points in time. This data included the results of questionnaires tapping attitude and behavior information, interview results, tallies from systematic observation, and standardized achievement test results. The data were analyzed by comparing the mean of an EXP school (one which had been offered incentives) to the mean of a matched CON school in the same city. The interpretation was restricted by problems inherent in the design and time schedule of the project. Although it is possible that results reflect factors other than the incentives offer, the achievement gains observed for students in the Parent-Teacher model were substantially greater than those of the control group. Differences in achievement gain between the Teacher-Only model and the control group were negligible. Other results from the analysis of the attitude and behavior data also are presented and discussed, but no simple patterns were evident in these results. (Author)

F I N A L R E P O R T

INCENTIVES IN EDUCATION PROJECT
IMPACT EVALUATION REPORT

Under Contract No.
OEC-0-71-4770

Submitted to

United States Office of Education
Department of Health, Education and Welfare
Washington, D.C.

Submitted by

The Planar Corporation
Washington, D.C.

October 24, 1972

ACKNOWLEDGMENTS

The Planar Corporation is pleased to submit this final impact evaluation report on the Incentives in Education demonstration project to the Office of Program Planning and Education.

The goals of field research in education are to improve the lives and prospects of us all. However self-evident and noble these goals may be, it must be explicitly recognized that they are rarely achieved without some real human costs. It is infrequent that people welcome the disruption of their lives through the imposition of an idea that others in remote places think deserves a test. This was particularly true of the idea and implementation of incentives in education. Impinging daily upon their lives, the evaluation of incentives had to be, indeed was, an intrusion that took its toll. That the principals, teachers, parents and children in these cities tolerated our presence with good will and understanding is to their great credit and has given us the incentive to interpret the information honestly and draw our conclusions wisely.

The logistics of field research and evaluation of this type requires the cooperation and patience of many people throughout the nation who are generally not to be found when the Final Report is submitted. In recognition of their dedication and the fact that this report was so dependent on their good work, The Planar Corporation field staff deserve special mention: Col. Robert Armstrong, U.S.A.F. (Ret.) in Jacksonville, Florida; Mrs. Marilyn Barbadora, Cincinnati, Ohio; Mrs. Carolyn Sanchez, Oakland, California; Rev. Aaron Jones and Mrs. Hilda Arevelos, San Antonio, Texas.

The design and execution of all of the analyses in this report were the responsibility of Dr. James Fennessey of The Johns Hopkins University. Working under Dr. Fennessey's supervison in this undertaking were Stephanie Freeman, Lawrence Howe, Paula Jaworski, Richard Mantovani and Carol Prevatt of the Planar staff. Everything from the layout to the printing of this report was under the management of Mrs. Sherry Van Dalen. Her professionalism and attention to every detail are deeply appreciated. Finally, a debt of gratitude is owed to the Center for the Social Organization of Schools,

and its School Organization Program, Johns Hopkins University, for its substantial contribution to the execution of the data analysis.

Mr. Edward Glassman of the Office of Planning, Budgeting and Evaluation was responsible for the design and implementation of this demonstration project and served as Project Officer for its duration. His understanding of the vicissitudes and ambiguities inherent in the logistics and analysis of field research made this project more pleasant than we had any right to expect.

The information contained in this report was provided pursuant to Contract OEC-0-71-4770 with the U.S. Office of Education, Department of Health, Education, and Welfare. The conclusions and recommendations reported herein are those of the Project Manager and Dr. Fennessey and should not be construed as representing the policy of any individual or agency of the United States Government.

Peter G. Briggs
Project Manager

Executive Summary

Introduction

The Office of Education designed and implemented a one-year project to demonstrate the feasibility of offering incentives to parents and teachers or to the teachers only. Separate process and impact evaluations of this demonstration were conducted under contract to the Office of Program Planning and Evaluation in the Office of Education. The process evaluation was conducted by Education Turnkey Systems, Inc. and can be found in two reports: Final Report, July 1, 1972, and Final Evaluation Report, July 31, 1972. This report presents the results of the impact evaluation conducted by The Planar Corporation.

Project Description

Two incentive models were employed in four medium-sized cities for one year in this project. In Cincinnati, Ohio, and Jacksonville, Florida, incentives of up to $600 per subject were offered to teachers, (Teacher Only model), depending entirely upon their classroom's performance on standardized Reading and Mathematics achievement tests. In Oakland, California, and San Antonio, Texas, incentives of up to $600 per subject were offered to teachers and incentives of up to $50 per subject were offered to parents, (Parent Teacher model) depending entirely upon the classroom's performance on standardized Reading and Mathematics achievement tests. For each classroom a predicted growth rate was calculated in both subjects, and the amount of payment to the parents and teachers depended upon the degree to which the classroom exceeded the predicted growth rate.

In each city there was one experimental school in which the incentives were offered and a matched control school. All schools in the project were elementary schools, Grades 1-6, with from 500 to 700 students, most of whom were low-income and underachieving. The project began in September, 1971 and ended in May, 1972.

Impact Evaluation Objectives

This project had extremely limited objectives: Only two out of a multitude of incentive models were implemented; only four cities with one building per city participated in the implementation; and only the feasibility of incentives was to be evaluated. Because so little is known about what happens when incentives are offered to parents and teachers in real world contexts, this evaluation covered a wide range of participants and outcomes, in both the experimental and control schools.

The major questions guiding the evaluation were:

> "What are the positive and negative results on student attitudes, behavior and achievement and parent and teacher attitudes and behavior of offering incentives in education, and what are the implications of these results for future Federal research?"

Findings

The results that are presented here must be accepted very cautiously by the reader. The project was delayed substantially by contractual problems, so that the teachers and parents did not sign their contracts until the school year was half over. Also, administrative delays prevented a full assessment of student, teacher and parent attitudes.

Student Achievement Results. In Reading, there was an educationally significant increase only in Oakland. In Cincinnati, there was an apparent decrease in learning rate.

In Mathematics, there were increases in learning in Oakland and San Antonio.

Student Attitudes Results. In Oakland there was a slight positive impact upon student attitudes. In San Antonio, there was an equally slight negative impact.

Student Behavior Results. No observed differences at any site.

Teacher Attitude Results. In Cincinnati, there was a slight negative impact. In Jacksonville, there was a slight positive impact. In San Antonio, there was a slight negative impact in which the teachers favored more traditional teaching techniques.

Teacher Behavior Results. In Cincinnati and in Jacksonville there is a trend toward more traditional classroom behaviors. In Oakland and San Antonio, there is a slight trend toward traditionalism.

Parent Attitude Results. In Cincinnati and Jacksonville there is a negative trend. In Oakland and San Antonio, there is a positive trend

Methodological Findings.

Student Achievement. The data were analyzed for the fan-spread phenomenon and the results of the analysis did not change the above findings. However, the achievement results are probably influenced by regression artifacts of an unknown magnitude.

Project Design. The small number of units and the nonrandom assignment of schools to treatment conditions limit the generalizations that can be drawn from this project.

Instrumentation. There was not enough time for the development, pretesting and clearance of the forms used in assessing attitudes.

Implementation. Unresolved questions of the funding authority hindered the full implementation of the incentive models and the evaluation.

Conclusions

Single Site Conclusions.

- In Cincinnati, the offer of incentives to teachers only had a negative impact.
- In Jacksonville, the offer of incentives to teachers only had a mixed impact, both positive and negative.
- In Oakland, the offer of incentives to parents and teachers had a positive impact.
- In San Antonio, the offer of incentives to parents and teachers had a slight positive impact.

Incentive Model Conclusions.

- The offer of incentives to teachers only had an overall negative impact.

- The offer of incentives to parents and teachers had an overall slightly positive impact.

Logistics and Management Conclusions.

- Single payment formulas, *e.g.*, one year's growth, are going to turn out to be unfair in most cases.

- School district data files are unreliable and are not a trustworthy source of primary or secondary evaluation material.

- Research projects become a part of the power conflicts and confusions in a school district.

- The traditional communications system between homes and schools in a district cannot bear the burdens imposed upon it by innovative projects.

- OE is not administratively structured to conduct adequately field projects of this type.

Recommendations

1. The Teacher Only incentives model should not be further researched or advocated.

2. The Parent Teacher model should be further researched.

3. Field research of incentives in education should provide for changes in the distribution of authority in school buildings and districts.

4. Field projects should have three phases, with each phase lasting at least one year.

5. As a means of gaining their responsible commitment to a field project initiated at the Federal level, school districts should share in funding the project.

6. Federal agencies that undertake field projects that are quasi-experimental in design and that necessitate a thorough evaluation must be internally structured so that the project officer has extensive authority. Fiscal, site selection, operational and evaluation decisions in projects of this type must be made swiftly and authoritatively.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

-Continued

LIST OF TABLES (Cont'd.)

- Continued

LIST OF TABLES (Cont'd.)

- Continued

LIST OF TABLES (Cont'd.)

- Continued

LIST OF TABLES (Cont'd.)

- Continued

LIST OF TABLES (Cont'd.)

- Continued

LIST OF TABLES (Cont'd.)

CHAPTER I

PROJECT DESCRIPTION

Background

Traditional educational methods do not seem to be effective for the low-income, underachieving student. Entering the public school system from a home of economic deprivation, this child is in a position of disadvantage from the very beginning. As he progresses through elementary and secondary school, the gap between the student of low-income background and other students tends to widen. The disparity is manifested primarily in terms of academic achievement, but there are major secondary effects in the areas of social adjustment, self-concept, career expectations, and level of aspiration. The child from a poverty background starts lower, learns less rapidly, and profits least educationally and economically from the traditional educational program offered by the public schools.

In personal, psychic terms the toll on the student is appalling. There is a Kafka-like atmosphere for the child. He plays a game the rules of which he does not understand, the purpose of which he is never able to comprehend, and the outcomes of which are almost certainly going to be unfavorable for him. It is small wonder that the low-income student soon withdraws, first spiritually, and then, when he can, bodily as well.

Social conscience and the simple economics of conserving human resources demand that methods be found to help the low-income underachiever to break out of this grim cycle and to perform academically on a par with his more affluent and fortunate contemporaries. Numerous educational innovations have been tried. At the Federal level, these innovations are relatively new; but they are increasing. Head Start and Home Start, aimed at lessening the initial disadvantage at a relatively early age represent one approach. Experimentation with new instructional techniques and media, as in the OEO performance contracting field experiment, is another. A third, which is the subject of the present study, involves the use of incentives with no control over program content. The basic assumption is that if incentives are offered

to teachers and parents, their motivation will increase and hence, the achievement of economically deprived students will increase.

Intrinsic rewards, mastery of a discipline or task, *etc.*, have traditionally been advocated by educators as the proper route toward more effective learning. However, this intrinsic reward system, upon closer examination, has been supported, indeed sustained, by an intricate system of rewards that are extrinsic to the act of learning. Extrinsic incentives are not new in the public schools. The educational system abounds with inducements and rewards to the students. There are "stars", grades, certificates, and other such tokens which are extrinsically valuable and may be exchanged for teacher and parent approval, peer esteem, or other psychic intangibles. Ultimately, academic success is tied to such long-term extrinsic incentives as a better job, wider opportunities for self-fulfillment, and a richer quality of life. This tacit assumption that education is directly related to upward mobility is shared by some teachers, parents, and the community as a whole. There is an ingrained tenet in the public school system that academic success leads, or should lead, to material reward.

For low-income children, however, the link between present achievement and future payoff is at best tenuous. Consequently, much of the recent experimentation with incentives in education has sought to use rewards systematically and to develop ways for making the reward structure explicit and more powerfully operative in the teaching-learning situation. Frequently this has involved supplying specific, immediate, and tangible rewards for classroom achievement. In a majority of the previous cases, the student has been the recipient; in the case of this project, the performance incentives were offered to teachers or to parents and teachers.

During the school year 1970-71, performance contracting appeared to be a promising means of systematically using incentives in school settings. The Office of Economic Opportunity (OEO) implemented a remedial education experiment in reading and mathematics. Federal support was given to 18 participating school districts for the subcontracting of remedial teaching to private educational technology companies. These contracts were of the performance contracting type, *i.e.*, the private companies were to be

remunerated in proportion to the reading and mathematics achievement gain of the students they taught throughout the year. To provide a parallel assessment of performance incentive contracting without using a private educational technology company, OEO contracted with two school districts to test the use of incentives in regular classrooms, where teachers were remunerated in proportion to the achievement gains of their students (Ray, *et al.*, 1972).

At the same time, the RAND Corporation, under contract to HEW, performed an evaluation of performance contracting at five sites throughout the country. The products of this evaluation were: an analysis of the performance contracting concept (Stucker and Hall, 1971); case studies of eight performance contracts (Carpenter, Hall, *et al.*, 1972); and a performance contracting guide (Hall, *et al.*, 1972). These RAND documents were designed to provide the educational community with a comprehensive and balanced appraisal of performance contracting.

Also in 1971, American Institutes of Research of Palo Alto, California, conducted a feasibility study which included a review of the literature on incentives, the development of incentive models for field testing, the identification of candidate school districts to participate in an experiment, and the identification of cost estimates. In addition, AIR also conducted a pilot study for six weeks in one school district. With a strong caution about overgeneralizing from such a limited experience, AIR concluded that objectives-based incentive techniques were effective in promoting student learning in basic skill areas. This pilot study also led to several recommendations for refining the incentive models and for enhancing the acceptance of the concept by parents and teachers (Jung, *et al.*, 1971; Lipe, *et al.*, 1971).

The basic proposition of all these recent experimental incentives programs can be stated quite simply. Incentives focus motivation. Motivation, in turn, influences learning. Hence, the systematic application of incentives in the teaching-learning situation should lead to superior student achievement. Beyond this theoretical assertion, however, there remained a number of practical questions. How are the incentives to be supplied? What form should they take? To whom should they be offered?

In the study which is the object of this report, attention was focused not upon direct incentives to students but upon incentives to other participants in the educational process, the teacher and the parent. In the case of the teachers, a traditional economic device was applied in a new context. A promise was made to pay a monetary bonus to the person charged with instructing the children if he or she produced student achievement in excess of the predicted rate of growth.

In the case of parents, a financial incentive was offered for their encouragement and support at home and at school to the extent that, as a result, their children attained a higher than predicted rate of achievement. Parent involvement is a recognized means of enhancing student performance, and the aim was to test the effectiveness of such a synergistic parent-teacher incentive model.

Although the project had some classical experimental features, it must be recognized that the intent was not to conduct a formal, structured field experiment. There were uncontrolled variations throughout the experimental design. There were some program variations, *i.e.*, departmentalization, types of books, *etc.*, between experimental and control schools within districts. There were also program differences across the four cities. No attempt was made to specify classroom procedures or teaching techniques. Through a small scale, field pilot project, the purpose was to take a close look at the consequences of teacher and parent-teacher incentive models, to learn more about the process by which they have their effect, and to test the reaction of the participants. In effect, this was a search for information and insights which could be used to make decisions on possible future experimentation in this area.

Objectives

The basic purpose of the project was to study the effects of incentives offered to teachers or to parents and teachers. There were two incentive models. In one, incentives were offered to teachers only in two cities. In the other, incentives were offered to parents and teachers in two cities. Three major outcomes were to be considered:

- Academic achievement of the students
- Behavior of the parents, students, and teachers
- Attitudes of the parents, students, and teachers.

Implicit within this general statement of objectives are a number of specific questions, which can be divided into process, product, and policy categories. This report addresses primarily the product questions. The final reports by Education Turnkey Systems, Inc. (1972a & b) deal with process questions. Both reports consider policy questions. The specific questions that were addressed in the reports were:

Product Questions

- What influence did incentives have on student achievement gains in reading and mathematics, as measured by norm referenced tests?
- What effects did incentives have on student attitudes toward reading and mathematics, the school in general, and themselves?
- Were there significant changes in academically related behavior of students attributable to the incentives program?
- Were there specific changes in teacher behavior as a result of the offering of incentives?
- Did the offering of incentives produce changes in the nature or degree of parental support and encouragement of student achievement?
- Did the incentives program produce changes in parental attitude toward their children's academic performance, the school system, or their children's teachers?

- With respect to all of the above, were there different effects attributable to the type of incentive model used?

Process Questions

- What factors influenced parent and/or teacher acceptance of, and participation in, the project?
- What influences can be attributed to local and specific differences among teachers, parents, administrators, or other involved groups?
- What obstacles were encountered in implementing the incentives program or in making it operate successfully?

Policy Questions

- What impact did the incentives program have on teacher and parent organizations, school administrators, community groups, educational decision makers, and local politicians?
- What changes should or could be made in the incentives models to increase student achievement in reading and mathematics?
- What changes should or could be made in the incentives models to make them more acceptable to professional, parental, and community groups or to eliminate specific undesirable features?
- What trade-offs were feasible in cases where pupil achievement gains and teacher or community acceptance are at odds?

Project Organization

Basic Features

The project was designed and implemented as a field demonstration under the sponsorship of the Office of Planning, Budgeting and Evaluation (OPBE) at OE. It was monitored and evaluated by two independent contractors. The Planar Corporation was the Testing and Analysis Contractor (TAC) and was charged with the responsibility for achievement testing, attitude assessment and other matters relating to product evaluation. Education Turnkey Systems, Inc., the Monitoring and Assistance Contractor (MAC), was responsible for process evaluation and for assisting OPBE and the local school districts in the administration of the project. This report deals only with TAC activities. Separate process evaluation reports have been issued by MAC (Education Turnkey Systems, Inc., (1972a & b). The following summary of the division of responsibilities between MAC and TAC is provided to clarify the organizational structure and the working relationships between the evaluation contractors and the participating school districts.

Two incentive models were employed in two cities each in this project. In two cities, incentives were offered only to teachers. Parents were informed of the project and were encouraged to support it, but they were not offered any incentives for their support. In the other two cities, both parents and teachers were offered incentives. In all four cities, control schools were selected, observed, monitored, and tested in exactly the same manner as the experimental school; however, incentives were not offered to the control schools. Hence, the only systematic differences were between experimental and control schools within cities and between incentive models in two pairs of cities.

The principal independent variable of the project was the offer of incentive payments. The secondary independent variable was the type of incentive model: Parent-Teacher or Teacher Only. In nondepartmentalized grades, classroom teachers were offered cash bonuses of up to $1200 ($600 for reading and $600 for mathematics) for increasing the mean achievement level of their students. In departmentalized grades, teachers were offered

up to $1200 divided by the number of classrooms he or she taught. In both situations, qualification for the bonus was determined on the basis of each classroom's achievement gain above a predicted level, as determined by that classroom's pre to post-test differences on norm referenced reading and mathematics tests. The bonus was not an all-or-nothing proposition. A formula was devised to allow graduated payment for actual gains above expected gains.

Incentive payments to the parent or guardian of students were offered on a similar, proportionate basis. Depending upon the pretest to post-test gain score of their child's classroom, the parent or guardian could receive up to $50 per child per subject. Thus, parents could obtain a maximum bonus of $100 for each of their children enrolled in the experimental school.

The target population for the project was children in Grades 1 through 6. A minimum of 80% were black children who were significantly below national norms in reading and mathematics. A high proportion of the children came from homes receiving some form of welfare assistance and from neighborhoods with low-income levels. These criteria were in keeping with the general aim of testing the incentive models as one means of addressing the chronic problem of the low-income underachieving student.

Monitoring and Assistance Responsibilities

The Monitoring and Assistance Contractor had the primary responsibility for making the initial protocol and working agreements with Federal, State and local officials and organizations. During the project, MAC maintained communication with the local school administrations, teachers, school board members, parents, and community representatives. MAC also monitored classroom activities in the experimental and control schools at each site and provided feedback to OPPE on the progress of the experiment. At the conclusion of the project, MAC prepared two reports on process evaluation. The major tasks performed by MAC in each phase of the project were:

Setup Phase

- o Help negotiate with school districts for their participation and conclude agreements for payment of incentives.
- o Prepare and distribute explanatory materials and conduct orientation sessions for project participants.
- o Assist TAC and OPPE in the development of the payment formula.
- o Develop a delivery system for the payment of earned incentives.
- o Prepare a monitoring plan and establish the mechanisms for project monitoring.

Monitor and Assist Phase

- o Monitor classes and activities at experimental and control schools.
- o Conduct continuing orientation and explanation of the project for teachers, students, parents, and other concerned groups.
- o Collect samples of test and curriculum materials in use in the participating schools and forward to TAC for the curriculum audit.
- o Report to OE on the progress of the project.

Incentive Distribution, Analysis and Reporting Phase

- o Help to deliver incentive payments to teachers and parents.
- o Assist in arrangements for satisfactory completion of the project (submittal of necessary data and documents, debriefing of key participants, certification of incentive payment delivery, preparation of news releases, etc.).
- o Conduct a cost analysis of the incentive models.
- o Prepare final reports describing project activities, explaining outcomes, and presenting conclusions on policy implications in the area of incentives.

Testing and Analysis Responsibilities

The following is a list of specific responsibilities of the Testing and Analysis Contractor. Activities carried out by TAC in support of these responsibilities and the results obtained are discussed in detail in subsequent sections of this report.

Setup Phase

- Collect and analyze school district data to support site and school selection.
- Design student, teacher, and parent questionnaires.
- Research existing standardized achievement tests and recommend appropriate tests to OPPE.
- Assist OPPE in clarifying project outcomes, developing performance measures and designing appropriate instruments.
- Design and maintain a biographic and background data bank for participating students.

Operational Phase

- Administer pre and post-tests in reading and mathematics.
- Design and develop a payment formula for teachers and parents based upon student pre-test scores in reading and mathematics.
- Administer project related instruments to students, teachers, and parents.

Analysis and Evaluation Phase

- Score and analyze all project related instruments.
- Analyze achievement scores in reading and mathematics.
- Analyze attendance data.
- Compute incentive payments for teachers and parents.
- Audit curriculum materials used in experimental and control schools to ensure that test items were not being taught.
- Prepare a final report evaluating project outcomes.

To carry out these responsibilities, TAC established a field organization at each of the four project sites. Data collection activities and administrative arrangements with the schools were under the direction of a TAC Field Coordinator, who was assisted by a group of Field Investigators. This working group conducted all on-site data collection for TAC and served as the nucleus of a much larger group of achievement test examiners hired locally to administer pre and post-testing of reading and mathematics achievement.

At each site, the local school administration appointed a District Project Director drawn from the administrative or teaching staff, who acted as the on-site coordinator and liaison channel between TAC personnel and the principals and teachers of the participating schools. Figure I-1 is a schematic representation of the project organization, showing the major lines of authority from OPBE down to the school level. It must be emphasized that Figure I-1 indicates only organization structure and that, in practice, a close working relationship existed among OPBE, local school district, TAC, and MAC personnel.

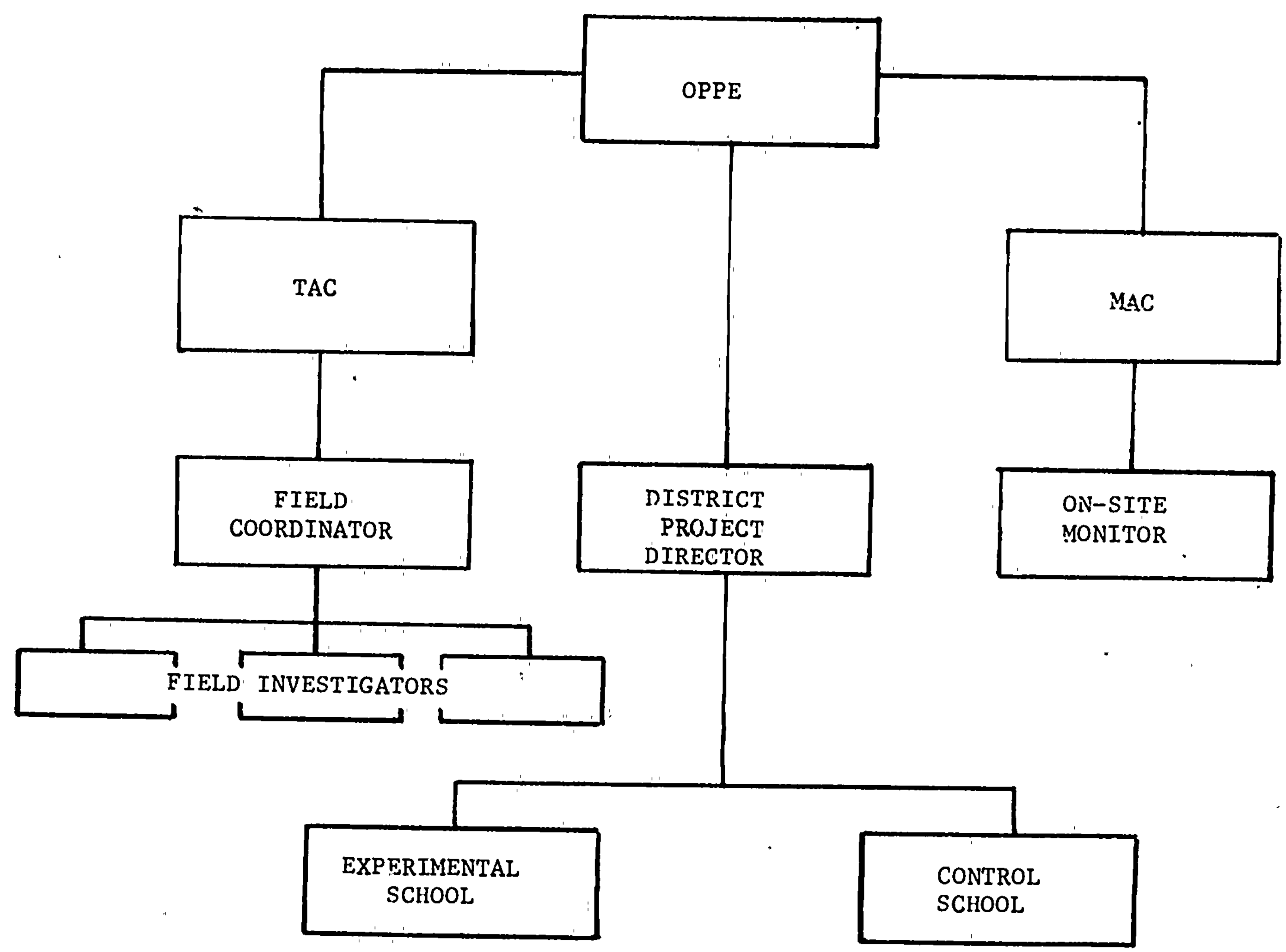


Figure I-1. Project Organization.

Project Procedures

Procedural Objectives

Basically, this project was a "hands off" field demonstration project. There was no intervention by MAC or TAC in the conduct of classroom activities or in the selection of instructional techniques employed by the teachers. Each teacher was completely free, within the limits of school policy, to use the method and/or materials of his or her choice to upgrade student performance in reading and mathematics. The absolute amount and frequency of testing, observing and visiting were the same in the experimental and control schools. Thus, the only experimental intervention was the incentive bonus offered in the experimental schools.

In educational projects of this sort it is impossible to obtain a laboratory-like control of the situation. There are always many uncontrolled sources of variance, and no claim is made to have conducted a pure experiment. However, the basic strategy of nonintervention was an attempt to give the incentives as free a play as possible while observing the incentives' impact upon other elements of the school and community situation. The experimental strategy and the purpose of the demonstration can, therefore, be summed up in a single question:

> If we take two comparable schools with a large proportion of low-income, underachieving students and introduce special incentives in one school what differences between the schools will emerge and what are the research implication of these results?

The procedures followed by TAC and MAC were conditioned by the basic policy of nonintervention. At the start of the project, MAC and TAC activities were limited to conducting orientation sessions for the school personnel and parents and to drawing up the necessary working agreements. During the school year, MAC provided assistance, as required, in project administration and monitored activities in the experimental and control schools. TAC served as the agency for data collection and administered all achievement test instruments. Naturally, these activities drew attention to MAC and TAC by students, teachers, parents, and the community at large; but the basic posture was as nonobtrusive as possible.

Monitoring and Assistance Procedures

The following sections describe the procedures to set up, administer, and monitor the project and to collect data for process evaluation.

School District and Teacher Agreements. MAC staff followed the Title III Guidelines and developed model proposals to be used for the Teacher-Only model and the Teacher-Parent model for the four school districts. Budgets were developed and discussed with each of the four sites to clarify total USOE commitment to the project and to identify the specific costs of personnel and services at the individual sites. MAC continued to provide management support and technical assistance during the proposal development process from 24 August 1971 through 20 December 1971 at which time each site received notification that its Title III grant award was forthcoming.

Participating sites were required to contract locally with the individual participating teachers and/or parents for the payment of incentives to be made at the end of the school year. MAC staff worked with the OE Project Officer to develop a model contract for teachers and a second contract for parents. The contracts were reviewed by administrative personnel at each site and revisions made to accommodate local administrative philosophies as much as possible without weakening the intent of the project. In several instances MAC staff worked directly with superintendents and school boards to arrive at an agreeable procedure by which the participating teachers could receive an incentive payment while under contract to the district to provide services as a teacher, and without violating the terms of the local association/union agreement.

This contract development process required seven revisions and modifications before reaching a final agreement which was acceptable to all concerned parties. Presentations were then made at all project sites to explain the details of the contracts and the nature of the formula to be used in determining the amount of incentives to be paid. This activity began on 29 July 1971 and continued through April 1972 at which time the last teacher contract was signed at the Duval County, Florida, site. Parent contract negotiation continued through May 1972.

Monitoring and Assistance. Continuous on-site monitoring and administrative assistance was the responsibility of MAC. The synopsis given below is intended to outline the role of MAC and to indicate the procedures for monitoring and assistance.

At the start of the school year, MAC prepared a "Standard Operating Procedures" manual for the school district Project Directors. The manual contained a description of the project, specification of procedures for administration and documentation, budgeting and accounting guidelines, and scheduling information. MAC also conducted a three-day workshop for Project Directors to orient them to the project, to review specific duties and responsibilities and to explain administrative procedures.

MAC established and maintained a system for collecting monitoring information and for serving as a clearinghouse for information about the conduct of the project. MAC also prepared and distributed explanatory material, public information releases, and background documents for use by news media.

Throughout the school year, the MAC representative at each site conducted a program of observations in the participating schools. This included observation of classroom activities, teacher and parent interviews, attendance at teacher and PTA meetings, and periodic meetings with school principals and district administrative personnel.

In concert with TAC, an audit of curriculum materials was performed by MAC. Items used in the classrooms were collected by MAC and forwarded to TAC to determine possible matching to specific items on the post-tests. The aim was to identify any potentially test biasing items, *i.e.*, classroom materials that were taken directly from the standardized achievement tests. Instances where such matching items appeared in classroom materials on other than a random or accidental basis were brought to the attention of OPPE.

Testing and Analysis Procedures

Data Collection, Testing and Analysis. TAC was charged with the major responsibility for data collection during the project. This included establishing and maintaining a basic biographic/demographic data file on

project participants, pre and post-testing of student achievement, classroom observations, and gathering information on attitudes and behaviors of students, parents, and teachers. With the exception of pre and post-testing, which required a large number of examiners hired and trained by TAC for that purpose, all data collection was carried out by the TAC Field Coordinator and a staff of four to six people working in cooperation with the school district Project Director.

A variety of data collection procedures were employed, including group tests, questionnaires, interviews, observations, and excerpts of school district records. Table I-1 is a summary of the methods used to collect data of each type for project participants. A brief description of these procedures is given below. A detailed discussion of the instruments, procedures, and actual schedules is provided in the following chapter under the heading of "Measurement Techniques".

TABLE I-1

SUMMARY OF DATA COLLECTION METHODS

		INFORMATION CLASS			
		Achievement	Attitude	Behavior	Demographic
Participant	Student	T	Q,I	Q,I,O,A,R	R
	Teacher		Q,I	Q,I,O	Q
	Parent		Q,I	Q,I	Q

Key: T = Achievement Test Q = Questionnaire I = Interview
O = Classroom Observation A = Attendance Record R = School Records

Achievement Tests. The Metropolitan Achievement Tests, 1970 edition, (MAT) were used as pre and post-tests in reading and mathematics and were administered to all project students in the control and experimental schools by school district substitute teachers trained and supervised by TAC. Homeroom teachers were present in the classroom but did not take part in the testing. Pretesting was carried out within the first three weeks after the start of school in each city, and post-testing was conducted three weeks prior to the end of the school year. Students entering both schools during the school year, up to the last month before post-testing, were also pretested. The data from these new students were analyzed separately to determine the impact of incentives on achievement rates over periods of time shorter than a school year.

Questionnaires. Questionnaires to elicit attitudes and to obtain self-reports of behavior were administered to all experimental and control participants twice during the year. Student questionnaires were administered in the classrooms by TAC examiners. At the same time, parent questionnaires were given to students to be taken home. Responses were mailed by the parents to the TAC project office in Washington, D.C.. The parent questionnaire dealt with attitudes, self-reports of behavior (specific assistance given to children, study facilities in the home, attendance at school meetings), and family demographic information. For both pre and post parent questionnaires, nonrespondents were randomly sampled and were sent a follow-up questionnaire. If the selected parents did not return the follow-up questionnaire, they were telephoned at home and requested to complete and mail it.

The teacher questionnaires were distributed through the schools and returned to the TAC project office in Washington by mail. In addition to attitudinal and behavioral information, the questionnaire contained items relating to teacher background, experience, and qualifications. All questionnaire responses from all participants were filed by a project-unique identification number that guaranteed the respondent's anonymity.

Interviews. All teachers in the experimental and control schools were interviewed by TAC field personnel twice during the year, in the fall and spring terms. In addition to questions pertaining to attitude and specific elements of teaching practice, free response questions were included

to elicit a broad range of opinions on incentives, the project itself, the students, the school, and the teaching profession. The free response items were analyzed by OE and MAC as part of the process evaluation. All teacher interviews were conducted in the school.

Fifty parents and 50 students per city were interviewed in the fall and spring by TAC field personnel. For each control or experimental school a stratified random sample of parents by homeroom was selected. To assure full and equitable coverage across all schools and sites, a 100% oversample was drawn. Parents were contacted and interviewed in their homes until a total of 25 interviews had been completed for each school. Care was taken to preserve an even distribution of interviews across homerooms. In effect, this amounted to two or three parent interviews per homeroom depending upon the size of the school. After the parents had been interviewed, arrangements were made through the teachers to interview the children of these parents during school hours. Thus, a matched set of parent-student responses was obtained. The post-test interviews were conducted with the same parents who had been interviewed in the fall. In those cases where the parents were not available for post-interviews, another name was selected from the same homeroom and then the parent's child was also interviewed. Free response items on the parent and student interviews were also analyzed by OE and MAC as part of the process evaluation.

Observations. TAC representatives conducted two series of classroom observations during the school year. Each classroom was visited for four 20-minute periods in November-December 1971 and again for four 20-minute periods in March-April 1972. Observers recorded teacher and student behavior according to a uniform schedule. Data were collected only on observed teacher and student behavior; no attempt was made by the observer to record attitudes or to draw inferences as to the meaning of classroom activities.

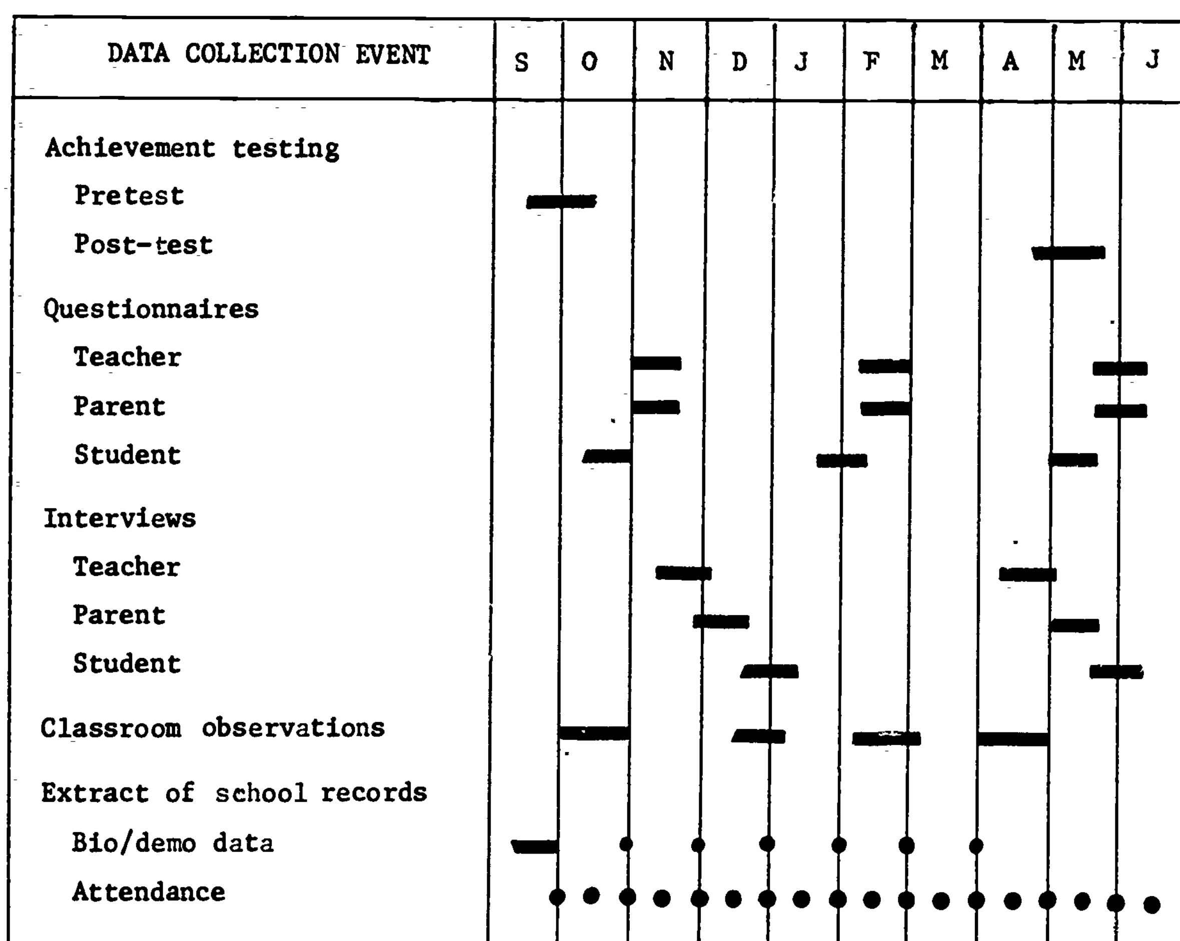
Extracts of School Records. School district Project Directors, with the assistance of local TAC field personnel, compiled a basic biographic/demographic record for each student in the project. This involved reviewing and excerpting school records to obtain such information as age, sex, address, siblings in the same school, years in the school, and so on. This was done

for each student in the control and experimental schools at the start of the program and for transfers when they entered either school. Supplementing this, a record was kept of attendance throughout the school year.

Table I-2 is a summary of data collection activities during the project. Each data collection event is listed according to the scheduled time of performance. However, it must be emphasized that the schedule shown here is only a nominal schedule and that certain deviations and delays were imposed by local conditions and unforeseen circumstances. The specific reasons for these departures from schedule will be taken up in the next chapter in the discussion of how each instrument was administered.

TABLE I-2

NOMINAL SCHEDULE OF TAC DATA COLLECTION



Payment Formula

The original intent of the project was to employ a simple incentive formula whereby the minimum achievement gain for incentive payments, the Base Gain Indicator (BGI), would be calculated by subtracting the mean pretest grade equivalent score for a given subject and grade level from the corresponding pretest score for the next higher grade level. That is,

$$BGI = \overline{X}_{n+1} - \overline{X}_{n}$$

Where BGI = Base Gain Indicator

$\overline{X}_n$ = mean pretest grade equivalent score for a given grade (n)

$\overline{X}_{n+1}$ = mean pretest grade equivalent score for the next higher grade (n + 1).

Thus, if the mean grade equivalent pretest score in reading for the second grade of a particular school were 1.8 and the corresponding third grade score were 2.6, second grade students in a given classroom would have to gain an average of 0.8 grade level on the post-test for their teacher to qualify for the minimum incentive payment.

As a result of conferences with local school officials and teachers during and after the pretesting, TAC and OPBE concluded that such a procedure might be inherently unfair to teachers with low achieving students in homogeneously grouped classrooms. To determine the extent to which homogeneous achievement grouping actually existed in the experimental schools, TAC conducted an analysis of the pretest scores of all reading and mathematics classes in the experimental schools. Using the chi-square test with the .01 level of significance as a criterion, the individual classroom groups were found to be significantly different in 37 of 48 cases (4 schools X 6 grade levels X 2 subjects). Thus, in over three-quarters (77%) of the classrooms, it could be assumed that students were stratified according to achievement level and that in a sizable proportion of these cases the concentration of low achievers would operate to the disadvantage of the teachers in attaining incentive payments.

Derivation of the Base Gain Indicator. For this reason, it was decided to derive a payment formula which would be sensitive to the initial achievement level (and, hence, the potential for gain) for each classroom. As a first step, the student population in each grade in each school was rank ordered twice; first, by MAT standard scores in reading, and then by MAT standard scores in mathematics. Standard scores were used instead of raw scores because there were as many as four different test levels used within a grade and raw scores are not comparable among test levels.

The next step was to divide each grade level in each experimental school into thirds (high, middle, and low achievers). The division point between one achievement level and the next was at that standard score which was closest to one-third of the total grade level for that school population. This procedure resulted in groups which were of slightly unequal size, but it assured that all students with the same standard score were placed in the same group.

Next, the means for each of the three groups in each subject were calculated, yielding six means for each grade level, three for reading and three for mathematics. Because there were instances of severe anomalies between the means for successive grade levels, it was necessary to calculate regression equations for the high, middle, and low reading and mathematics groups across the six grade levels. By then fitting a regression curve to the means, the base gain indicators were smoothed and no longer subject to inordinate variations from one grade level to the next. Several different regression equations were tried, but the one which produced the best fit across all sites, Grades 1 through 5, and achievement groups was a second order polynomial:

$$L_p = a + bG + cG^2$$

Where L_p = predicted level of student achievement expressed as a grade equivalent score

G = present grade level

a,b, and c = empirically derived coefficients.

Table I-3 presents the predicted grade equivalent scores derived from the regression equation. Because this regression formula predicted post-test scores for Grade 6 that were significantly greater than could be reasonably expected, a linear regression equation was used for that grade level.

As a final step, a base gain indicator was derived. The BGI is the least amount of gain that a classroom had to achieve for the teacher to obtain a bonus payment. It was calculated for a given grade level by subtracting the actual classroom pretest score for a given achievement group from the predicted pretest score for the same achievement group (high, medium, or low) at the next higher grade level.

Thus:

$$BGI = L_{P_{n+1}} - L_{P_n}$$

Where BGI = Base Gain Indicator

L_{P_n} = predicted level of student achievement (in grade equivalent score) for a given grade (n)

$L_{P_{n+1}}$ = predicted level of student achievement (in grade equivalent score) for the next higher grade (n + 1).

Thus, for reading in Grade 2, for example, there would be three BGIs, one each for the high, middle, and low achievement groups. Table I-4 lists the BGIs for each site, grade, achievement group, and subject.

To determine which BGI to apply to a given classroom teacher, the average pretest score for the individual classroom was computed for each subject. Then it was determined which achievement interval (high, middle, or low) that classroom fell within. The BGI for that achievement group was then assigned to the classroom.

Teacher Payment Formula. A graduated scale of four payment increments was established for the teachers. To obtain a minimum payment for one subject (25% of the $600 maximum, or $150), the teacher of a given self-contained class had to raise the students' mean achievement level, as measured by the post-test, by an amount equal to the BGI. For each 0.1 grade equivalent above the BGI, the teacher obtained an additional $150. The maximum bonus of $600 per subject was thus attained if the mean

(Test resumes on page I-28)

TABLE I-3

PREDICTED MEAN GRADE EQUIVALENT SCORES FOR EXPERIMENTAL GROUP STUDENTS

Reading				Arithmetic			
Oakland	High	Medium	Low	Oakland	High	Medium	Low
1	1.67	1.41	1.27	1	1.36	1.18	1.02
2	2.12	1.59	1.34	2	1.98	1.48	1.13
3	2.86	1.99	1.54	3	2.73	1.97	1.40
4	3.89	2.65	1.87	4	3.74	2.67	1.84
5	5.21	3.39	2.32	5	4.96	3.58	2.45
6	6.79	4.41	2.90	6	6.38	4.68	3.23
7	8.41	5.61	3.60	7	8.02	5.92	4.17
Cincinnati				Cincinnati			
1	1.43	1.28	1.23	1	1.16	1.05	0.94
2	2.13	1.69	1.42	2	2.11	1.45	1.19
3	2.88	2.12	1.65	3	2.99	1.96	1.52
4	3.68	2.57	1.90	4	3.82	2.57	1.93
5	4.54	3.04	2.19	5	4.57	3.29	2.42
6	5.45	3.53	2.50	6	5.27	4.10	3.00
7	6.40	4.03	2.83	7	5.90	5.02	3.66
San Antonio				San Antonio			
1	1.41	1.27	1.20	1	1.28	1.08	0.94
2	2.12	1.64	1.40	2	1.99	1.47	1.15
3	2.82	2.01	1.61	3	2.77	1.95	1.43
4	3.52	2.39	1.83	4	3.63	2.52	1.79
5	4.23	2.78	2.06	5	4.55	3.19	2.21
6	4.93	3.17	2.30	6	5.54	3.94	2.70
7	6.33	3.57	2.60	7	6.60	4.79	3.06
Jacksonville				Jacksonville			
1	1.39	1.21	1.17	1	1.19	0.96	0.87
2	1.69	1.43	1.28	2	1.52	1.19	0.99
3	2.21	1.69	1.41	3	2.12	1.60	1.21
4	2.92	2.01	1.56	4	2.98	2.18	1.53
5	3.83	2.37	1.73	5	4.11	2.93	1.96
6	4.95	2.78	1.91	6	5.50	3.85	2.48
7	6.27	3.24	2.12	7	7.15	4.95	3.09

TABLE I-4

BASE GAIN INDICATORS--CINCINNATI

GRADE LEVEL	ARITHMETIC ACHIEVEMENT LEVEL		
	Low	Medium	High
1	.25	.41	.95
2	.33	.51	.89
3	.41	.61	.82
4	.49	.71	.76
5	.58	.81	.70
6	.66	.91	.63

GRADE LEVEL	READING ACHIEVEMENT LEVEL		
	Low	Medium	High
1	.20	.41	.70
2	.23	.43	.75
3	.25	.45	.80
4	.28	.47	.86
5	.31	.49	.91
6	.34	.51	.96

TABLE I-4 (Continued)

BASE GAIN INDICATORS--JACKSONVILLE

GRADE LEVEL	ARITHMETIC ACHIEVEMENT LEVEL		
	Low	Medium	High
1	.12	.23	.34
2	.22	.40	.60
3	.32	.58	.86
4	.42	.75	1.13
5	.52	.92	1.39
6	.62	1.10	1.65

GRADE LEVEL	READING ACHIEVEMENT LEVEL		
	Low	Medium	High
1	.11	.22	.31
2	.13	.26	.51
3	.15	.31	.71
4	.17	.36	.92
5	.19	.41	1.12
6	.21	.46	1.32

TABLE I-4 (Continued)

BASE GAIN INDICATORS--OAKLAND

GRADE LEVEL	ARITHMETIC ACHIEVEMENT LEVEL		
	Low	Medium	High
1	.10	.30	.58
2	.27	.50	.79
3	.44	.70	1.00
4	.61	.90	1.21
5	.76	1.10	1.42
6	.94	1.30	1.64

GRADE LEVEL	READING ACHIEVEMENT LEVEL		
	Low	Medium	High
1	.07	.18	.45
2	.20	.39	.74
3	.33	.60	1.03
4	.45	.81	1.32
5	.58	1.01	1.60
6	.71	1.22	1.89

TABLE I-4 (Continued)

BASE GAIN INDICATORS--SAN ANTONIO

GRADE LEVEL	ARITHMETIC ACHIEVEMENT LEVEL Low	Medium	High
1	.21	.39	.71
2	.28	.48	.78
3	.35	.57	.85
4	.42	.66	.92
5	.49	.75	.99
6	.56	.84	1.10

GRADE LEVEL	READING ACHIEVEMENT LEVEL Low	Medium	High
1	.20	.37	.70
2	.21	.37	.70
3	.22	.38	.70
4	.23	.39	.70
5	.24	.39	.70
6	.25	.40	.70

achievement gain of the class was 0.3 grade equivalent or more above the BGI. If the class failed to attain the BGI, the teacher received no bonus. Table I-5 shows the incentive payment schedule in relation to achievement gains.

TABLE I-5

TEACHER INCENTIVE PAYMENT SCHEDULE
FOR SELF-CONTAINED CLASSROOMS

Mean Post-Test Achievement Gain	Teacher Incentive Payment
<BGI	\$0
BGI	\$150 ÷ total number of classrooms taught that subject
BGI + 0.1 G.E.	\$300 ÷ total number of classrooms taught that subject
BGI + 0.2 G.E.	\$450 ÷ total number of classrooms taught that subject
BGI + 0.3 G.E.	\$600 ÷ total number of classrooms taught that subject

It should be remembered that the teacher was eligible for an incentive payment in both reading and mathematics. If, for example, the mean class gain in reading were BGI + 0.1 and in mathematics BGI + 0.2, the teacher received a bonus of \$750. In departmentalized grade levels, the maximum incentive per classroom was \$1200 divided by the number of classrooms taught by that teacher. The payment schedule was then adjusted to provide payment levels within that maximum.

Parent Payment Formula. The schedule of incentive payments to parents was similar to that of teachers. Parent payment was dependent upon the performance of their child's classroom, not the child's individual achievement. The maximum incentive bonus to parents was \$50 per subject,

payable in increments of $12.50 according to the schedule in Table I-6.

TABLE I-6

PARENT INCENTIVE PAYMENT SCHEDULE
FOR READING OR MATHEMATICS

Child's Classroom Achievement Gain	Parent Incentive Payment
> BGI	$0
BGI	$12.50
BGI + 0.1 G.E.	$25.00
BGI + 0.2 G.E.	$37.50
BGI + 0.3 G.E.	$50.00

Incentives were paid to parents for each child in the experimental school. A parent with two children, with each child's classroom making a gain of 0.3 grade equivalent above the BGI in both reading and mathematics, would therefore qualify for a bonus of $200.

Project Sites

The incentives experiment was conducted at schools in four cities: Cincinnati, Ohio; Jacksonville (Duval County), Florida; Oakland, California; and San Antonio, Texas. These four sites were selected by OE from a list of school districts identified by OE as potential participants. The site and school selection process, which was a joint activity by MAC and TAC under the direction of OPBE, is described below.

Site Selection

At the outset of the project, a list of candidate school districts was drawn up by OE. The following criteria were used to determine initial eligibility:

- The school district had to be in a medium or large city, population between 250,000 and 2,000,000.
- The school district had to include large absolute numbers of black children, who represented a high percentage of the school population.
- The school district had to include large numbers of children from low-income families, on both an absolute and proportional basis.
- The school district had to be in a state which had unobligated ESEA Title III, Section 306 funds (Commissioner's Discretionary Fund) available at the start of the project.

The criterion of medium-city size was finally adopted for two reasons. First, it was consistent with the overall project aim of testing incentives models in an urban setting. Second, by excluding major metropolitan areas in favor of medium-sized cities, it was possible to avoid administrative and logistical problems inherent in very large school districts and, at the same time, to assure that there would be a sufficient number and concentration of minority, poor students. The second and third criteria follow logically from the basic objective of demonstrating incentives as a possible approach to helping underachieving, low-income students. The fourth

criterion (availability of Title III funds) was included for administrative reasons. The funds to pay school district administrative costs for the project and to pay for incentives earned by teachers and parents were drawn from ESEA Title III, Section 306 (Commissioner's Discretionary Fund).

More than 50 school districts were determined by OE to be eligible according to these five criteria. Telephone and mail inquiries were made to these eligible districts by MAC under the direction of OPPE to ascertain their willingness to participate in the project. Eleven districts indicated that they wished to be considered and visits were made to all of these candidate sites by OPBE, MAC and TAC representatives. The final selection of the four participating districts was on the basis of: the degree of the district's desire to participate in the project; the ability of the district to obtain cooperation from the local teacher organization; the likelihood of favorable community attitudes; and the availability of schools meeting specific criteria to serve as either experimental or control schools.

Table I-7 is a summary of the characteristics of the participating local school districts in terms of the objective selection criteria. The table also indicates which incentive model was selected by each district. A tentative decision on the incentive model was made at the time the site was selected. The final decision was based upon the preference of the school district administration and the school board. This requirement that local preferences prevail entailed a change from the original assignment of a model in Cincinnati. The school administrators in that city had originally accepted the Parent-Teacher model; however, prior to the pretesting, the school board decided to accept the Teacher-Only model. Fortunately, the San Antonio school board was willing to change from the Teacher-Only to the Parent-Teacher model, thereby preserving the basic design of the project.

TABLE I-7

CHARACTERISTICS OF PARTICIPATING SCHOOL DISTRICTS

SELECTION CRITERIA	PROJECT SITES			
	Cincinnati	Jackson-ville	Oakland	San Antonio
Population[1]	452,524	528,865	361,561	654,153
Elementary School Enrollment[2]	50,206	54,986	35,950	41,087
Number Minority Children[2]	22,365[3]	20,316[3]	26,251[4]	33,396[4]
Percent Minority Children[2]	44.5[3]	30.7[3]	73.0[4]	81.3[4]
Number Poor Children[2,5]	11,388	18,090[6]	10,472[6]	14,003
Percent Poor Children[2,5]	22.7	32.9	29.1	34.1
Title III Available	Yes	Yes	Yes	Yes
Incentive Model	Teacher Only	Teacher Only	Parent Teacher	Parent Teacher

1. Based on 1970 census figures.
2. Based on school enrollment 1970-71.
3. Black children only.
4. Includes Black, Spanish American, Oriental, American Indian, and other nonwhites.
5. According to Title I poverty criteria.
6. Proportionate estimate based on total school enrollment.

School Selection

The selection by the OE Project Officer of individual schools eligible to participate in the project was based upon the following considerations:

- The school was participating in an ESEA Title I program or, if not participating, was eligible for Title I funds.
- The school had at least 80% black children from poor homes.

- The school had Grades 1 through 6.
- The school was not participating in Federally supported instructional or remedial programs other than ESEA Title I.
- The school had a demonstrated record of underachievement, as measured by recent standardized tests in reading and mathematics.

The schools in each district were ranked by TAC in terms of these five criteria and then grouped into pairs, matched as closely as possible to ensure comparability between experimental and control groups. Each pair of candidate schools was then screened by a group consisting of OPPE, MAC, TAC, and district administration members to make a final choice. Among the characteristics which led to selection of the pair of schools to participate in the program were:

- Achievement test scores below the other pairs
- A high percentage of low-income students
- The likelihood of support by the school principals
- The likely acceptance of the program by the teaching staff, school parents, and the immediate communities.

Whenever possible, as part of the final selection process, interviews were conducted with principals and teacher representatives. Specific points of discussion were: the willingness of the principal and teachers to cooperate in the project, and assurances by the school district that the teaching staff from the previous year would be retained, except for those who already had an approved commitment for transfer. The experimental school was chosen on the basis of its having lower academic performance and more children from poverty homes than the control school.

The experimental and control schools finally selected for participation in each city are listed below:

<u>Cincinnati</u>

Garfield Elementary School (experimental)
Cummins Elementary School (control)

Jacksonville

Livingston Elementary School (experimental)

Carver Elementary School (control)

Oakland

Longfellow Elementary School (experimental)

Santa Fe Elementary School (control)

San Antonio

Washington Elementary School (experimental)

Tynan Elementary School (control)

Table I-8 on the following page is a summary of the student population characteristics in the experimental and control schools. Note that two control schools are listed in Oakland. A group of teachers in the control school originally selected, Washington Elementary School, decided at the start of the school year to withdraw from the project. A suitable replacement, Santa Fe Elementary School, was arranged just prior to the scheduled start of pretesting. Both schools are listed in Table I-8 for the sake of comparison, but it should be understood that only data from Santa Fe were used in determining project outcomes in Oakland.

Table I-9 contains student achievement test results in reading and mathematics for 1970-71, the year before the start of the project. Results are tabulated under the grade level of the students in the year in which tested. Scores are reported both in grade equivalent form (in the column titled G.E.) and in terms of grade level decrement (in the column titled DEC). The grade level decrement is an approximate index of how far the schools were behind national norms. Thus, if the mean achievement test score for fourth grade students tested in October was 3.5, the decrement is listed as -0.7, which represents the difference between what they should have scored on the second month of the fourth year (4.2) and what they actually scored (3.5). Grades in which tests had not been administered by the school districts are denoted by the entry "NG".

TABLE I-8

CHARACTERISTICS OF PARTICIPATING SCHOOLS

SCHOOL	No.[2] Students	No. Black	% Minority	No.[2,3] Poor	%[2,3] Poor	TITLE I Elig.	TITLE I Recip.
Cincinnati							
Garfield (E)	612	576	94.1	351	61.0	Y	Y
Cummins (C)	649	636	98.0	230	42.5	Y	Y
Jacksonville							
Livingston (E)	889	887	99.8	504	56.7	Y	Y
Carver (C)	1019	1018	99.9	342	33.6	Y	Y
Oakland							
Longfellow (E)	608	592	97.4	258	42.4	Y	N
Washington (C)[1]	611	591	96.7	310	50.7	Y	N
Santa Fe (C)	856	836	97.7	503	58.8	Y	N
San Antonio							
Washington (E)	654	620	94.8	236	36.1	Y	Y
Tynan (C)	429	418	97.4	181	42.2	Y	Y

1. Withdrew at start of project, replaced by Santa Fe.
2. 1970-71 enrollment figures.
3. Family meets Title I poverty criteria.

TABLE I-9

STUDENT ACHIEVEMENT LEVELS, 1970-71 SCHOOL YEAR

R E A D I N G

SITE/SCHOOL	GRADE 1		GRADE 2		GRADE 3		GRADE 4		GRADE 5		GRADE 6	
	G.E.[1]	DEC.[2]	G.E.	DEC.	G.E.	DEC.	G.E.	DEC.	G.E.	DEC.	G.E.	DEC.
Oakland, Longfellow (E)[3]	1.7	-0.2	2.5	-0.4	2.2	-1.7	NG[4]	----	4.1	-1.2	5.0	-1.3
Oakland, Santa Fe (C)	1.7	-0.2	2.7	-0.2	2.9	-1.0	NG	----	3.9	-1.4	4.7	-1.6
Oakland, Washington (C)	1.7	-0.2	3.0	+0.2	2.9	-1.0	NG	----	3.7	-1.6	5.0	-1.3
Jacksonville, Livingston (E)[5]	NG	----	NG	----	NG	----	2.8	-2.0	NG	----	5.0	-1.8
Jacksonville, Carver (C)	NG	----	NG	----	NG	----	2.9	-1.9	NG	----	4.3	-2.5
Cincinnati, Garfield (E)[6]	NG	----	NG	----	3.03	----	NG	----	NG	----	3.28	----
Cincinnati, Cummins (C)	NG	----	NG	----	3.48	----	NG	----	NG	----	3.68	----
San Antonio, Washington (E)[7]	NG	----	NG	----	NG	----	3.0	-1.2	NG	----	NG	----
San Antonio, Tynan (C)	NG	----	NG	----	NG	----	3.1	-1.1	NG	----	NG	----

NOTES:

1. G.E. = Grade Equivalent Score on nationally normed achievement test.
2. DEC. = Actual Grade Equivalent Score minus Expected Grade Equivalent Score.
3. Oakland achievement tests administered in May, Grades 1-4, and November, Grades 5-6.
4. NG = Tests not given for that particular grade/subject.
5. Jacksonville achievement tests administered in April, all grades.
6. Cincinnati scores reported are stanine score for 50th percentile.
7. San Antonio achievement tests administered in May, all grades.

TABLE I-9 (Continued)

STUDENT ACHIEVEMENT LEVELS, 1970-71 SCHOOL YEAR

M A T H E M A T I C S

SITE/SCHOOL	GRADE 1		GRADE 2		GRADE 3		GRADE 4		GRADE 5		GRADE 6	
	G.E.	DEC.	G.E.	DEC.	G.E.	DEC.	G.E.	DEC.	G.E.	DEC.	G.E.	DEC.
Oakland, Longfellow (E)	NG	----	NG	----	NG	----	NG	----	4.2	-1.1	4.6	-1.7
Oakland, Santa Fe (C)	NG	----	NG	----	NG	----	NG	----	3.8	-1.5	4.8	-1.5
Oakland, Washington (C)	NG	----	NG	----	NG	----	NG	----	3.7	-1.6	4.8	-1.5
Jacksonville, Livingston (E)	NG	----	NG	----	NG	----	3.0	-1.8	NG	----	4.7	-2.1
Jacksonville, Carver (C)	NG	----	NG	----	NG	----	3.4	-1.4	NG	----	4.6	-2.2
Cincinnati, Garfield (E)	NG	----	NG	----	2.26	----	NG	----	NG	----	3.50	----
Cincinnati, Cummins (C)	NG	----	NG	----	2.23	----	NG	----	NG	----	3.79	----
San Antonio, Washington (E)	NG	----	NG	----	NG	----	3.0	-1.2	NG	----	NG	----
San Antonio, Tynan (C)	NG	----	NG	----	NG	----	3.2	-1.0	NG	----	NG	----

CHAPTER II

METHODOLOGY

Introduction

Although the project was a field demonstration and not a formal experiment, the selection or development of measurement instruments and their administration were guided by formally stated evaluation objectives. A process was followed in which desired outcomes were stated as hypotheses, items were developed (or instruments selected in the case of achievement outcomes) and then the items were assigned to the instruments.

There were two objectives in this activity. The first objective was to ensure that all parties involved in the design of the project had an opportunity jointly to present their views about the outcomes. Too often in the design of large scale projects or field evaluations, because there are many uncontrolled sources of variance and because several groups participate in the design, outcomes are not treated systematically or comprehensively during the planning phase. Consequently, there are gaps in the data and many unanswered questions at the conclusion of these projects. The instrument development process used in this project was designed to avoid these inadequacies. The second objective was to ensure that items or questions developed to measure the same hypotheses were assigned to different instruments. Because interviews, questionnaires, and observations were to be used by both MAC and TAC, it was desirable that the reliability of the measurement be increased by spreading these items out over the several different types of data gathering activities employed by the two contractors. There were ten different data collection instruments developed and used by TAC and MAC in this project. Ranging from inventory checklists to open-ended interview protocols, these instruments provided a unique opportunity to distribute items related to the same variable among a wide variety of instruments.

Obviously, the systematic development or selection of instruments is the necessary but not sufficient first step for an adequate evaluation of any program. The specificity of the questions, the sensitivity of the data collection techniques, the appropriateness of the analysis and the intelligence of the interpretation are critical intervening variables between this development process and the final report. The variable and instrument development process followed in this project was important only in its power to facilitate the performance of subsequent tasks and decisions.

Achievement Measurement

Instrument Selection

Achievement Specifications. Four specifications were established by OE pertaining to the measurement and evaluation of student achievement. Two of these were outcome specifications; one was procedural; and one was an evaluation strategy specification. The first outcome specification was that only Reading and Mathematics achievement were to be measured. Schools teach many skills and attitudes, but Reading and Mathematics were selected for attention because of the discouraging pattern of low-income school results in these two areas. Furthermore, because this project was as much a demonstration of the logistical feasibility of offering incentives as it was an experiment, Reading and Mathematics were defined in the narrowest sense of the two terms. Consequently, for example, student language arts and spelling skills were not to be considered. It was believed that by focusing upon a narrow range of skills, the issue of the practical feasibility of incentives could be clearly addressed in this field demonstration.

Procedurally, because the tests were to be used to determine payments as well as to evaluate the impact of the offer of incentives, it was necessary to guarantee that there was no opportunity to "teach the test". Since the time of the allegations about teaching the test in the Texarkana, Arkansas, and Providence, Rhode Island performance contracts, any project that involves payment for results, whether it is performance contracting or teaching incentives, must account for this factor. Multiple testing (one test for payment and one test for evaluation) with both tests disguised or masked has been the usual complete solution to this problem. However, its cost in terms of time, student patience and teacher good will are formidable practical barriers. OE was aware of these conflicting considerations and had specified that the achievement assessment in this project be designed to resolve them.

Because this project was conceived and administered at the Federal level, the paramount evaluation objective was that the achievement results be measured in clear and understandable terms. The Federal Government has many audiences, and it was believed that the achievement results of this project should be based upon instruments and analyses acceptable to both educational researchers and practitioners. This requirement meant that the achievement tests meet high psychometric standards and yet not be so arcane as to be uninterpretable by those who are not research specialists. In short, it was important that the educational and statistical meanings of this project be considered in the selection .nd subsequent analysis of the achievement tests.

<u>Test Selection</u>. The Metropolitan Achievement Test Batteries, 1970 Edition, (MAT), published by Harcourt Brace and Jovanovich, Inc., were selected for administration after a careful examination of a variety of measurement instruments and other factors. In view of the many valid criticisms of performance contracting instrumentation and analysis, the decision to select the MAT series was not taken lightly. In arriving at these decisions, the following factors were considered:

- Because of the populations in this project, it was decided that test batteries should be assigned individually to each student on the basis of the best available information about his or her current Reading achievement level. This meant that the achievement test levels would be mixed within grade levels and that these mixes would vary between experimental and control schools within sites and among the four sites. Therefore, the achievement tests had to have a score scale that would allow comparisons across test levels for evaluation and payment gains analysis. The results of the MAT, Primer through Intermediate Batteries, are expressed for all batteries and all forms of those batteries on a single, common scale. This made the MAT uniquely suited to the assignment of tests on an individual basis.

- Because this project did not impose any materials or methods upon the teachers, the achievement test content had to be appropriate to the wide range of curriculum materials that would be used throughout the project. In addition, because of the rapid curricula changes in the past few years, it was necessary that the achievement tests have recently developed editions so that the content validity would be high. The MAT had undergone an extensive content revision in 1970.

- Because the participants in the project were low-income, inner-city students, it was essential that the norming population be representative of the demographic trends in the nation. With the urbanization of the population, and with low-income, inner-city children increasing as a percentage of public elementary students, the achievement tests had to be normed on a representative and recent sample. The MAT had been restandardized in 1970.

- The tests had to meet acceptable psychometric validity and reliability standards. The MAT has separate fall and spring forms and norms.

- This project was seen as part of a continuum of Federal research and experimentation in education. The OEO performance contracting experiment and the OE Anchor Test study were important elements in that continuum, and it was desirable that the results of this project be comparable to those of at least these other two major efforts. The MAT was used for the evaluation of the OEO experiment and was the "anchor test" in the OE Anchor Test study.

- Finally, there were a myriad of practical considerations. Because the project involved widely separate and disparate sites, there had to be at least two alternative forms. The format and administrative procedures had to be appropriate to the populations of interest. And the scoring facilities had to be convenient. The MAT series was highly satisfactory in all these respects.

Test Description

Norm Data. The Primer, Primary I, Primary II, Elementary and Intermediate levels of the Metropolitan Achievement Tests were selected as the test series for this project. This series was one of the most recently revised in terms of content and norm data. The salient norm data for the MAT standardization group and for the project populations are presented in Table II-1. As can be seen from the table, the MAT standardization population is quite representative of the national population; approximately 44,000 students in Grades 1-6 participated in the MAT standardization.

Although the recent MAT norming was representative of the national population and certainly comprehensive, it must be emphasized that the student population in this project was a subpopulation, not identical with the total norm group. The figures for the project population presented in Table II-1 are based upon questionnaires sent to all of the parents in March and April of 1972. Out of a possible return of 4500 questionnaires, only 1329, or 31%, were received. Thus, it is clear that generalizations from the respondents to the total population should be made with great caution. From an examination of the two populations in Table II-1, it can be seen that there are vast differences between the project population and the norm populations. Because of the urbanism, low income, and minority representation found in the project population, there are obvious limitations to the appropriateness of the tests' content. The children in the project come from a highly distinctive background. The content of the tests and, perhaps more importantly, the testing procedures themselves, were not designed with these children in mind.

TABLE II-1

COMPARISON OF STUDENT POPULATIONS,
PROJECT (1972), MAT STANDARDIZATION (1970),
SELECTED DEMOGRAPHIC CHARACTERISTICS

DEMOGRAPHIC CHARACTERISTIC	PROJECT POPULATION[1]	MAT POPULATION[2]
City size		
- 250,000	100%	21%
- 25,000-249,000		21%
- 2,500-24,999		30%
- Less than 2,500		27%
Family income	$3639	$5550
Minority Group		
- Black	95%	13%
- Spanish-speaking	3%	4%
- Oriental	.1%	4%
- American Indian	----	.3%
Years of schooling	12.2	10.7
H.S. graduates, parent	44%	41%
Unemployed	26%	5.1%

Sources:
1. Incentives Project Parent Questionnaire.
2. Harcourt Brace Jovanovich, Inc., Report No. 7, June 1971.

Despite these limitations of content and the reliance upon only one type of test-taking skill, the use of standardized tests such as the MAT is justifiable for evaluation and payment in a project of this kind. The content of the Reading and Math test particularly reflects the kinds of skills which the public schools strive to develop in all pupils. These skills, moreover, are demonstrably essential to success in later schooling

and in most occupations. Also, there is really no feasible alternative to the use of such tests for a project of this kind. The standardized tests are relatively sophisticated and adequate instruments for assessing the performance of children in terms of a generally accepted curriculum content and with the use of widespread teaching practices.. To have attempted to develop and use some other form of test instrument would have been an endeavor beyond the scope of this or any other demonstration project.

Finally, one other point which makes the use of these tests defensible in this project is that the results obtained from the children in the incentives schools are _not_ being compared with national norms, but with the results obtained from quite similar children in quite similar schools. Thus, any lack of appropriateness of the content of the tests, or the testing procedures, should not contribute any bias to the evaluation of the impact of the offer of incentives.

Test Content. Tables II-2 and II-3 present the skill content for Reading and Mathematics, respectively, of the five test batteries in the Metropolitan Achievement Test series. The subtest scores are added together to produce a Total Reading or Total Mathematics score. Subtests that did not contribute to these scores were not administered.

Test Reliabilities. The publisher's estimates of test reliability were derived using Saupe's estimate of the Kuder-Richardson Formula 20 and are presented in Table II-4. Because the publisher's standardization population was so dissimilar to this project's population, reliability estimates computed by the Battelle Memorial Institute (Ray, *et al.*, 1972) have been included in Table II-4. The Battelle reliabilities were derived from the MAT 1970 used for the evaluation of the OEO performance contracting experiment conducted during the 1970-71 school year. The OEO experiment involved low-income, underachieving students in Grades 1-3 and 7-9 with a higher than average minority representation. The Battelle reliability estimates were derived using a more conservative formula (KR-21) than the publisher's (KR-20). It is encouraging to note

TABLE II-2

MAT SKILL CONTENT--READING

TEST LEVEL	APPROP. FOR GRADES	SUBTEST	DESCRIPTION	NO. ITEMS
Primer	K.6-1.5	Reading	Letters - recognition of capital and lower case letters.	11
			Words - marking words that identify common pictures.	16
			Sentences - reading sentences that describe a picture. (A few items for the more able pupils, used in post-test only.)	5
Primary I	1.5-2.4	Word Knowledge	A sampling of primary level words presented in the familiar word-picture association format.	35
		Reading	Thirteen sentences in picture-sentences format, a few riddles, and short paragraphs.	42
Primary II	2.5-3.4	Word Knowledge	A continuation of the word-picture association format and introduction of short written sentences to be completed.	40
		Reading	A few sentences in the sentence-picture format. Reading selections ranging in length from a few sentences to fully developed paragraphs.	44
Elemen.	3.5-4.9	Word Knowledge	A sampling of vocabulary covering synonyms, classifications.	50
		Reading	Easy to difficult paragraphs with questions on main ideas, drawing inferences, getting literal meaning, use of vocabulary in context.	45
Intermediate	5.0-6.9	Word Knowledge	Covers a variety of words from social studies, science, humanities, general experience. Also antonyms, synonyms, classification.	50
		Reading	Selection ranging from simple paragraphs to fully developed topics with several paragraphs. Questions call for identification of main ideas, literal meaning, inferred meaning, and special meaning of words as determined by context.	45

TABLE II-3

MAT SKILL CONTENT--MATHEMATICS

TEST LEVEL	APPROP. FOR GRADES	SUBTEST	CONTENT	NO. ITEMS
Primer	K.6-1.5	Numbers	Concepts - basic ideas about time, money, geometry, counting, sets.	20
			Computation - addition and subtraction algorithms with one-digit numbers.	14
Primary I	1.5-2.4	Computation	Free response items with addition and subtraction (no regrouping). Horizontal and vertical notation covered.	27
		Concepts	Counting, reading and writing numerals, place value, and other basic concepts. Simple dictated problems involving addition and subtraction.	35
Primary II	2.5-3.4	Computation	A variety of addition and subtraction examples some involving regrouping. A few multiplication examples.	33
		Concepts	Measurement, place value, laws and properties and other modern concepts.	40
		Problem Solving	Seventeen teacher-directed items and eighteen pupil-read items on everyday numerical application.	35
Elemen.	3.5-4.9	Computation	Addition, subtraction, multiplication with and without regrouping. Introduction of division and fractions.	40
		Concepts	Modern coverage of laws and properties, sets, measurements.	40
		Problem Solving	A variety of problems involving arithmetic operations. Simple chart reading. Selecting appropriate number sentences.	35
Intermediate	5.0-6.9	Computation	Operating with natural numbers, decimals, fractions, percents; simple equations and negatives at advanced level.	40
		Concepts	Factors, primes, inequalities, estimation, laws and properties, concepts of fractions and decimals, other modern topics.	40
		Problem Solving	Application of geometry and measurement, use of charts and number sentences, multiple step problems, plus broad coverage of arithmetic operations.	35

TABLE II-4

RELIABILITY COEFFICIENTS FOR THE
METROPOLITAN ACHIEVEMENT TESTS

TEST LEVEL	FORM	TOTAL READING Publisher[1] (KR20)	TOTAL READING Battelle[2] (KR21)	TOTAL ARITHMETIC Publisher (KR20)	TOTAL ARITHMETIC Battelle (KR21)
Primer	F	.93	---	.96	---
	H	.93	---	.94	---
Primary I	G	.96	.98	.96	.97
	F	.96	.92	.94	.88
Primary II	G	.96	.96	.95	.94
	F	.96	.94	.95	.98
Elementary	G	.96	---	.96	---
	F	.96	---	.96	---
Intermediate	G	.96	.94	.95	.95
	F	.96	.93	.96	.93

Sources: 1. Harcourt Brace Jovanovich, Inc., Report No. 10, June, 1971.
2. Battelle Memorial Institute, March, 1972. (Battelle used the Stanford Early School Achievement Test and the California Achievement Test, Level I, for the first grade. The MAT Elementary Test was not used by Battelle because Grades 4 and 5 were not in the OEO experiment.)

that the two reliability estimates are very similar, despite the differences between the norm and the OEO populations. Hence, because the OEO and the OE and the incentives project populations were quite similar, it is reasonable to conclude that the publisher's reliability coefficients are applicable to this analysis.

Achievement Test Administration

Objectives. There were three objectives in the design and implementation of the achievement test plan. These objectives were:

- To reduce the error variance attributable to the students' taking test levels that are inappropriate to their achievement levels
- To reduce the error variance due to variations in testing procedures and to test circumstances that are not conducive to maximum performance
- To protect the integrity of the payment plan and the evaluation.

Procedures.

Test Level Assignment. Each student in the project was assigned to a test level that was appropriate to his or her current Reading achievement evidence. Based upon a review of the schools' achievement patterns over the past years, it was decided that the test publisher's recommended level assignments were not appropriate to this population. Collectively, the students in this project were achieving at about one grade level below their grade placement. Therefore, it was felt that a test battery that was one level below what the publisher recommended would be more appropriate. At the same time, it was recognized that there was a considerable amount of variation within the grade levels and that the most appropriate procedure would be to administer several test levels within a grade level. Because this was a logistically complex testing procedure, it was decided not to assign individually both Reading and Mathematics test levels. Table II-5 presents the intercorrelations between the Total Reading and Total Mathematics scores for the MAT fall standardization. Following a review of these intercorrelations and discussions with the test publisher, it was decided that the student's Reading level would be adequate for estimating the Mathematics test level.

TABLE II-5

INTERCORRELATIONS, TOTAL READING AND TOTAL MATH, MAT, BY TEST LEVEL

PRIMER	PRIMARY I	PRIMARY II	ELEMEN.	INTERMEDIATE
.66	.68	.74	.79	.76

Source: Harcourt Brace Jovanovich, Inc., Report No. 11, June 1971.

To assign a pretest level to a student, his or her prior Reading achievement test score was used where it was available from school records. Only test scores that had been derived from a spring 1971 administration were used, so that in no site were the data more than four months old. For those students for whom there were no test data, the teachers estimated in a grade equivalent format, *e.g.,* 1.4, 5.3, *etc.,* the student's Reading achievement level as of September 1971. After the teachers did this, the class lists were returned to TAC and a test level assignment was then made. No students were assigned a test level higher than the publisher's recommendations, and in only a few cases were test levels assigned more than two grade levels below the recommended grade level.

The same test level was used for the post-testing. It was recognized that this procedure might mean that some students' scores would be inaccurate because the scale score would not be high enough to reflect their true growth. However, this limiting factor applied to less than 5% of the students, evenly distributed between experimental and control schools. Furthermore, the possibility that item-analysis research might be done on these test scores made a test level change undesirable, because the subtests are not directly comparable across test levels. For these reasons, the pretest and post-test levels remained the same. Table II-6 gives the reader an indication of the distribution of test levels by grade level and building throughout the total project. As can be seen from the table, the test level assignments were generally one to two grade

levels below the publisher's recommended level, with the number of below grade level assignments increasing as the grade level increases. For example, the Primary II level, recommended by the publisher for administration in Grades 2 and 3, was not given to any second grade students; whereas 30% of all the Primary II tests were given to fifth grade students. This is congruent with the pattern of past achievement results revealed when the schools were selected (in which the schools fell increasingly behind the national norm as the grade placement increased).

TABLE II-6

INCENTIVES IN EDUCATION
TEST LEVEL ASSIGNMENT BY GRADE AND BUILDING, ALL CITIES

TEST LEVEL & PUB.'s RECOMM.	GRADE LEVEL											
	1		2		3		4		5		6	
	E	C	E	C	E	C	E	C	E	C	E	C
Primer (K.6-1.5)	364	310	187	189	122	58						
Primary I (1.5-2.4)			141	124	123	132	135	149	13	1		
Primary II (2.5-3.4)					88	161	140	131	172	136	91	87
Elemen. (3.5-4.9)					11	1	125	106	176	183	158	195
Intermediate (5.0-6.9)									21	46	128	153

Test Administration. TAC had the responsibility for selecting, hiring, training, and supervising test administrators for both the experimental and control schools in the four sites. All testing was under the supervision of a TAC staff professional. Local substitute teachers were hired and given one full day of training in achievement test administration, with the following two full days used to administer the achievement tests to all eligible students. (Students whom the school district had classified

as retarded for any reason were not a part of the project.) This all took place on a Wednesday through a Friday at each site. The following Monday and Tuesday were devoted to administering makeup tests to students who had been absent during the regular test administration. These procedures were followed for both the pretesting and post-testing.

Table II-7 puts the pre and post-testing into the perspective of the total school year at each site. The project test sequence was set up so that each school district received the pretest three weeks after school had opened. The three week delay was chosen because it was believed that the students needed some time to become accustomed to the work-discipline skills that are so necessary in taking these achievement tests. The post-tests were scheduled about three weeks before the end of school so that the students would not be restless from their anticipation of summer vacation and the end of school. The times of pre and post-testing were coincidental with the dates on which the publisher had normed the tests--October for the fall standardization and May for the spring standardization.

TABLE II-7

SCHOOL YEAR AND PRE AND POST ACHIEVEMENT TESTS

SITE	SCHOOL FIRST DAY	PRETEST FIRST DAY	POST-TEST LAST DAY	SCHOOL LAST DAY	NO. DAYS, SCHOOL	NO. DAY, PROJ.	% PROJ. YR.
San Antonio	8/30/71	9/23/71	5/4/72	5/31/72	180	130	72%
Jacksonville	9/7/71	9/30,71	5/4/72	6/8/72	180	140	78%
Cincinnati	9/8/71	9/30/71	5/12/72	6/14/72	180	139	77%
Oakland							
Longfellow (E)	9/13/71	10/7/71	5/19/72	6/15/72	180	130	72%
Santa Fe (C)	9/13/71	10/14/71[1]	5/19/72	6/15/72	180	125	69%

Notes: 1. The original Oakland control school withdrew from the project just prior to pretesting, which necessitated a one-week delay to locate another control school and assign test levels to students.

Two full days were set aside for both the pre and post-test administrations. Because students in any given classroom would be taking different test levels, the students had to be regrouped according to the level of the test they were taking. Within these groupings, the range of grade levels was generally not greater than two, so that the older students would not suffer a loss of self-esteem from taking tests with the younger students. Table II-8 displays the two-day test program, by test level and by subtest. The amount of time available for the limited testing was more than sufficient to ensure that the tests were given in a relaxed manner, that there were ample breaks between the subtests, and that the dictated tests at the lower levels could be accurately administered. The regular classroom teachers remained in the test rooms, while the Planar examiners administered the tests. The classroom teachers assisted in the logistics of the tests and in establishing an atmosphere that was conducive to test-taking.

<u>Results</u>. Despite the logistical problems inherent in multilevel testing, the pre and post administrations were accomplished with no problems that were likely to be an important source of error variance. For the major evaluation comparisons, by sites and by building, the test administration circumstances were essentially the same within the test administrations and between the pre and post-tests. For the payment computations, conducted for the experimental schools only, there was only one exception. A first grade classroom in Jacksonville was not pretested because the children had not had an opportunity to attend preschool and consequently had not learned how to hold a pencil.

Obviously, any testing program that is sponsored by the Federal Government and upon which payment and professional self-image are contingent is going to be accompanied by some tension within the schools. However, the school faculties were highly cooperative and the achievement testing conditions were more than adequate.

TABLE II-8

TEST ADMINISTRATION SCHEDULE

TEST LEVEL	PROJECT GRADES	TEST TIMES, MINUTES	
		First Day	Second Day
Primer	1-3		
Reading		20	
Numbers			25
Primary I	2-5		
Word Knowledge		15	
Reading		30	
Math			30
Primary II	3-6		
Word Knowledge		18	
Reading		30	
Math Comp.		18	
Math Con.			20
Probl. Solv.			25
Elementary	3-6		
Word Knowledge		15	
Reading		25	
Math Comp.		35	
Math Con.			25
Probl. Solv.			30
Intermediate	5-6		
Word Knowledge		15	
Reading		25	
Math Comp.		35	
Math Con.			25
Probl. Solv.			25

Attitude and Behavior Measurement

Objectives

Self-reports of attitudes and behaviors, observations of behavior, and pre-existing school building records of behavior were utilized in this project to cover as many outcomes as could reasonably be anticipated. In addition to the interest in student outcomes, teachers and parents were also included in the identification and development of outcome variables. Finally, in the original project design, the measurement of attitudes and behavior was to be conducted three times during the year in both the experimental and control schools. This was planned so that trends could be established which might provide some evidence of rates of impact over time and the sequences of change among the several outcome variables.

Instrument Development

The process outlined in the introduction to this chapter was followed in the development of attitudinal and behavioral measures. At each of the four sites, all instruments developed for this project were reviewed and critiqued by school officials. In no case was any original outcome variable or item dropped from the instruments, although several changes in format or method of administration were incorporated based upon the school officials' recommendations. The general considerations underlying each of the instruments follow. Appendix A contains examples of the instruments developed by TAC for the project.

To clarify questions that will undoubtedly arise from the discussion of and administration of various instruments, it is necessary to point out that the original design for the assessment of attitudinal and behavioral outcomes was not followed because of delays encountered in the OE and OMB forms-clearance process. The teacher, student and parent questionnaires were submitted by TAC in October, 1971, and did not receive final approval until the end of February, 1972. The changes that resulted from this five-month review are worthy of note:

- There were no changes to the substance or procedures for the student questionnaire.

- The parent questionnaire response scale was changed from a five- to a three-point scale because OE felt that a five-position scale was too complex for the project population. Follow-up procedures were modified slightly.

- There were no changes to the substance or procedures for the teacher questionnaire.

The original design of the project called for pre, interim, and post administrations of the instruments. Because of the forms-clearance delays, the questionnaires were administered only twice, in March and in May, an interval that hardly qualifies as a pre and post-test design.

<u>Student Questionnaire</u>. The student questionnaire was designed to measure a student's attitudes toward the teacher, the school, himself, and his peers. In addition, it attempted to measure self-esteem, feelings of control over immediate and remote situations, and future expectations and aspirations. There were 43 items in the instrument and there was no fixed time limit for the group administration. Students in Grades 1-3 were administered a practice questionnaire the day before the actual administration. For these grades, the TAC examiner dictated the questions to the students. The pages of the questionnaire were color coded with one question per page. In Grades 4-6, the questionnaire was self-paced. Twelve items were in the smiling-face, neutral-face and sad-face response format, with the responses for each question in the same order of presentation. The remaining 31 questions were in the balloon-child and flag-child response format. For each item, the balloon child appeared on the left and the flag child on the right. However, in some cases, the balloon child was associated with the more desirable answer choice, and in other cases with the less desirable answer choice. This was done as a means of eliminating some of the response halo effects that might otherwise operate. In both types of questions, the student was asked to choose the face or figure most like himself.

All students in the project were twice administered the questionnaire in the classrooms. The first administration was delayed by the required forms-clearance procedures from October, 1971 to March, 1972. The second administration was in May, 1972.

Student Interview. The student interview contained 15 items, five of which were self-reports on behavior and ten of which were assessment of the student's attitudes. The interviews were conducted by a TAC field representative in the school buildings and the student's responses were coded by the field representative during the interview on a five-point scale, ranging from positive to negative. The first part of the question was designed to give the student an opportunity to elaborate on his feelings toward the topic under consideration. This elaboration was noted by the field representative; and if a conflict appeared between the student's discussion and his subsequent response, the field representative probed further to ensure that the student's attitudes were accurately recorded. The questions covered attitudes and behaviors toward school in general, teachers, peers, reading and arithmetic, and future expectations.

Between 40 and 50 students per site (20-25 per school building) were interviewed twice during the year. The first interviews were conducted during the months of November and December, 1971, and the second interviews were conducted in April, 1972. Unless the student had left the school district, the same students were interviewed both times. Students were randomly selected by building, grade level, and classroom.

Parent Questionnaire. The parent questionnaire contained 45 items. Part I of the instrument contained 18 items designed to measure on a three-point scale attitudes toward their child's school, teacher, and academic progress. Twenty items were self-reports of the parent's or child's behaviors in areas thought to be relevant to academic performance. Part II of the instrument contained seven questions about the family's demographic characteristics. Parts I and II were administered in March, 1972; and Part I only was administered in May, 1972.

All parents in the project received the questionnaire which was distributed at the end of the school day and taken home by the students. The parents were asked to complete the questionnaire and mail it directly to TAC in a prestamped and addressed envelope. Three weeks after the questionnaires had been distributed, the nonrespondents were randomly sampled by site, building and grade level. No less than 50 parents per building were selected for the follow-up and were directly mailed a second questionnaire. Ten parents per building were sampled from those who did not return the follow-up questionnaire. They were telephoned at their homes by the TAC field representatives and encouraged to return the questionnaire.

Parent Interview. The parent interview contained 20 items, 13 of which were self-reports on behavior and seven of which were attitude assessments. The interviews were conducted in the parent's homes, and all responses were recorded by the TAC field interviewer. As with the student interview, there were two parts to each question, the first part open-ended and the second part closed and direct. The response to the second part was recorded. The questions on attitudes were recorded on a five-point scale ranging from positive to negative. The questions on behaviors were also recorded on a five-point scale with frequency counts assigned to each interval to eliminate subjectivity in recording the response.

Between 40 and 50 parents per site were interviewed in November, 1971, and in April, 1972. They were the parents of the students who had been selected for the student interviews, and they were interviewed in both the fall and spring.

Teacher Questionnaire. The teacher questionnaire contained 39 items. Part I of the instrument consisted of 25 questions about the teacher's attitudes toward the students, the parents, the profession, and the project; and three questions about teaching behavior. Part II contained 11 demographic and background questions. Parts I and II were administered to all teachers in March, 1972; and Part I only was administered in June, 1972.

The questionnaires were distributed in the schools, and the teachers returned them directly to TAC in a prestamped and addressed envelope. Nonrespondents were not followed up.

Teacher Interview. The teacher interview contained 14 items, seven of which measured attitude and seven of which measured behavior. All interviews were conducted in the schools, and responses were recorded by the TAC interviewer during the interview. Again, like the student and parent interviews, there were two parts to each question, one part open-ended and the other part closed. The open-ended responses were used by the interviewer to "break the ice" with the teacher and also to determine the accuracy of the categorization of the cloded response question. The responses to the attitude questions were recorded on a five-point scale ranging from positive to negative. The behavior questions were also recorded on a five-point scale with frequency counts assigned to each interval on the scale. All teachers were interviewed, once in November, 1971, and a second time in May, 1972.

Classroom Observation. The classroom observation instruments were designed for the systematic and direct categorizing and recording of student and teacher behavior. The instrument had 14 types of teacher behavior, *e.g.*, testing, lecturing, praising, *etc.*; and four context categories--one student, small group, large group and total class. There were also 14 categories of student behavior, *e.g.*, reading, listening, criticizing, *etc.*; and four context categories--alone, small group, large group and total class.

Classroom observation data were collected on every project teacher. There were two rounds of classroom observations, the first round in November and December, 1971, and the second round in April, 1972. During each round, a teacher's classroom was observed four separate times, about 15 to 20 minutes per observation. The TAC field representative's sequence of observations was as follows:

- Allow time for the class to settle down and for the lesson to begin.

- Begin the observations with the teacher. Observe, categorize, and record the behavior every ten seconds for three minutes.
- Then observe the students in the same manner for the next three-minute cycle. Each ten-second observation was taken on a different student.
- Go back to the teacher for another three-minute cycle and then to the students for a final three-minute cycle.

The four observations were added together each round to give pre and post observation scores.

School Records. Usually, a sad tale can be told whenever an attempt is made to use school document files for evaluation purposes; this project is no exception. Unproductive but substantial efforts were made in the following areas:

1. Attendance. The original plan called for collecting attendance data by individual student. This was found to be too time consuming; consequently, attendance data were aggregated and collected by homeroom group.

2. Tardiness. All schools in the project were supposed to keep tardiness data. In fact, only about 60% of the teachers kept reliable records of this behavior and there was no uniformity even within buildings. This effort was abandoned.

3. Library Use. Although all schools had libraries, some schools, because of the nature of their operation, were unable to keep any useful information on library use, and others could not even operate the library because of reductions in the budget. This effort was abandoned.

4. Other Information. In those districts that used the MAT in their achievement test program, the possibility of using post-test scores for a trend analysis was explored. This turned out to be unfeasible because mobility from year to year reduced the number of students to an unworkable sample.

4. Other Information (continued).

The use of school grades was explored, but different grading standards among the districts and within schools in one district, (Grades 1-3 used satisfactory-unsatisfactory whereas Grades 4-6 used A through F) precluded the use of this variable.

Disciplinary referrals were examined, but it was apparent that neither the referring teachers nor the principals' offices in any building maintained these data.

In summary, school building records were a disappointing source of reliable information. If the project had been planned to extend over several years, an information system might have been established that would have been as useful to the administration of the school as to the evaluation of the project. Obviously, the benefits of trend analyses are far greater than a one-shot experimental design; and the combination of a decent design and execution with historical information would have provided rich and reliable data not commonly found in educational evaluations. It should be noted, however, that the schools are really not at fault for these data management shortcomings. Schools do maintain reliable data when those data are useful to them. For example, attendance and enrollment data that generally determine State funding allocations were maintained and collected in all buildings in the project. The failure of the schools to keep information beyond mere attendance data is an outgrowth of the kinds of questions that are asked and the incentives offered for maintaining the information. A one-year, one-shot demonstration project is not a sufficient incentive to the schools to rearrange their normal recordkeeping procedures.

CHAPTER III

EVALUATION

Evaluation Methodology

Evaluation Strategy

The strategy that has been adopted for this evaluation analysis has several elements. Each of these elements was chosen in light of the policy objectives of the project, the constraints that actually operated in the field settings, and the need for rapid processing of the data and delivery of a final evaluation report.

The strategy included the collection of a considerable variety of data--achievement test scores, questionnaires, interviews, systematic observation, and analysis of school records--from the students, the teachers, and the parents. Much of this data was collected at more than one point in time. The purpose in collecting this widely diverse data was to provide maximum possible opportunity for <u>any</u> impacts of the incentives program, not just student academic achievement, to manifest themselves and be noted. The complementary efforts of MAC and TAC along these lines, in which MAC carried on formal observations and interviews at each site, and TAC collected the more quantitative data, were quite extensive. An effort was made to "leave no stone unturned" in detecting the potential consequences of the incentives program, and identifying the conditions which facilitate or inhibit its impact.

A second feature of the evaluation strategy was an effort to organize the research questions specifically around the policy issues described in documents provided by OE. To this end, a large number of hypotheses were formulated, involving considerable detail. These hypotheses then were organized according to a classification scheme that allowed them to be considered in groups. The classification scheme has three main components:

1. The particular comparisons that are to be made (*e.g.*, Parent-Teacher Model versus Teacher-Only Model)
2. The persons to whom the hypotheses refer (*e.g.*, students)
3. The specific dependent variable being examined (*e.g.*, reading achievement).

In general, all hypotheses were formulated so that the more favorable outcome was predicted for the treated group, and for the Parent-Teacher model (hereafter PT) in comparison with the Teacher-Only model (hereafter TO). In this introductory section, the first of the classification components--the particular comparisons to be made--will be discussed in some detail. Discussion of the other aspects of the hypotheses classification will be deferred until the sections on results. It should be noted here, however, that the number of specific hypotheses being examined is quite large. Thus, unless some organizing scheme and some summary hypotheses are used, there can be a real problem of losing sight of the forest amid the trees.

A third feature of the evaluation strategy was the decision to concentrate primary attention on the practical importance of whatever impacts might be found. In other words, the report will not merely report the statistical significance of differences in the outcome variables. In all too much educational research, including some very recent work, the investigators report that the treatment produced a statistically significant result in favor of their hypotheses, but give scant attention to the more important point that the actual size of the difference found is so small as to make it negligible from a practical point of view. Thus, the strategy adopted here was that statistical significance tests should be used primarily for one purpose: to verify that a difference which seems to be practically important (is numerically large in comparison to what was hypothesized) is not due merely to chance. Of course, in any analysis where a large number of comparisons are being carried out, some of these comparisons will exceed the critical value of the statistical test simply by chance. This is unavoidable. The appropriate response to it is to use caution in interpreting individual statistically "significant" results, especially if they are not part of a substantively coherent pattern.

The statistics used in tests of significance do have a second kind of application, however; and that application will be employed in this report. Such statistics as, for example, the value of student's t, or the point biserial correlation coefficient from a least squares fit, can serve as summary descriptions in common and familiar terms of the size of a particular difference relative to some other difference. Some of these summary statistics will be used as a means of presenting the results of comparisons. Substantive patterns in the results will also be identified as an additional means of deciding whether the separate results reflect some real and dependable impact of the incentives models.

A fourth element in the evaluation strategy was made possible by the design of this project, and made desirable by the nature of the audience for this report. That element is an attempt to rely wherever possible on simple and direct statistical methods. This will make the meaning of the analysis clearer to many readers. It will also minimize the number of debatable assumptions that are required for the application of the analytical method. It is a general principle that the more elaborate and sophisticated the statistical method, the more complex and questionable are the assumptions required for its legitimate application. In this report, statistics which are direct and straightforward will be applied wherever possible. The possible slight loss in "efficiency" incurred by this decision is not a problem, because it will not bias the estimates of the size of the treatment impact. At the worst, these simpler statistics will have slightly less precision of estimate, or will be slightly conservative as to the question of statistical significance. This is not a problem, given the policy-related purpose of this study.

Analysis Decisions

These general elements of the evaluation strategy prompted several decisions early in the evaluation analysis. Some of these are discussed here; others will be mentioned as they arise in the detailed presentation of results. The first decision concerned the score format to be used in analyzing the achievement test data. Because the analysis involved only

internal comparisons among the students involved in the project, there was no actual need for using scores which are referrable to national norms. However, it was thought desirable to relate the results of the achievement test analysis to the typical scores obtained by average students, so that the practical magnitude of the treatment impact can be judged more readily by the reader.

Three score forms were seriously considered for use in the achievement test analysis. These were: raw score, grade equivalent score, and test publisher's standard score. Note that the "standard score" does not refer to a score obtained by subtracting the sample mean and dividing by the sample standard deviation (as the term is often used in statistics). The main reason for not using the raw score was that the multilevel test administration, (in which each child in a class was tested with the test level deemed most appropriate to his personal achievement level at the time of testing) made any summary of the raw scores not comparable within or among the four sites. If raw scores had been used, an equating formula developed on only the project data would have to be created. Such a development would have been possible in principle, but in practice it would have been difficult and undependable, given the relatively small size of the sample taking each test level. For these reasons, the raw score analysis was rejected.

This left the standard score format and the grade equivalent format as possibilities. In the end, it was decided to adopt the standard score format, for reasons which will be explained below. First, however, it is necessary to describe briefly the standard scores and their relationship to grade equivalents. The standard scores are the score form used by the test publishers as the basic benchmark from which the grade equivalents are calculated. That is, the raw scores on each form and level were calibrated by the equipercentile procedure, and the resulting common scale, after some further smoothing, were used as the test publisher's standard scores (Orr, 1972). Thus, the standard scores are, in a certain sense, more fundamental than the grade equivalents because there is one less smoothing operation involved in their calculation.

Another argument leading to a preference for the standard scores is that they have less of what Donald Campbell (Educational Testing Service, Inc., 1970) has called "fan spread". That is, the standard deviations of standard scores for different test levels are more nearly constant than the standard deviations of the grade equivalents. Table III-1 shows the standard deviations for the norm groups on several of the test levels at several grade levels. As can be seen from the table, there is considerably more spread between the standard deviations in the grade equivalent scores (a range from 0.61 to 2.07 for Total Reading, and from 0.71 to 1.77 for Total Math) than there is for the standard scores (a range of 10.0 to 14.7 for Total Reading and from 10.4 to 12.6 for Total Math).

TABLE III-1

STANDARD DEVIATIONS OF METROPOLITAN ACHIEVEMENT TESTS DERIVED SCORES FOR PUBLISHER'S NORM GROUP

TEST BATTERY	GRADE LEVEL AT WHICH ADMINIS.	STANDARD DEVIATIONS			
		TOT READ		TOT MATH	
		G.E.	STAN	G.E.	STAN
Primary I	1.7	0.61	10.0	0.71	12.6
Primary I	2.1	0.77	10.9	0.74	12.1
Primary II	2.7	1.01	10.9	0.86	11.1
Primary II	3.1	1.17	11.6	0.92	11.4
Elementary	3.7	1.38	13.0	1.11	12.0
Elementary	4.1	1.61	14.0	1.19	12.0
Elementary	4.7	1.67	14.3	1.35	12.1
Intermediate	5.1	1.83	13.5	1.29	10.4
Intermediate	5.7	1.88	13.0	1.51	12.2
Intermediate	6.1	2.07	14.7	1.68	12.1
Intermediate	6.7	1.97	13.5	1.77	12.7

Source: Harcourt Brace Jovanovich, Inc., Report No. 8, June 1971.

In short, although either of these score forms will be affected by the fan-spread phenomenon, that effect will be less severe for the standard scores than for the grade equivalent scores.

An additional advantage of the standard scores in that the procedure used to create them and link the different levels of the test serves to justify the claim that the standard scores reflect equal intervals on the underlying achievement dimension being measured. In other words, the standard scores can justifiably be treated as interval scores. This is, in fact, their defining characteristic, and the one which makes it legitimate to carry out averaging and other arithmetic operations on them.

Also, the standard scores are more precise than the grade equivalents. Between the grade equivalent scores of 1.1 and 6.8, there are 57 units on the grade equivalent scale (as far as the actual data coded on the tape supplied by the scoring service is concerned). For the Total Reading Subtest, however, there are 70.5 standard score units (=85.5 - 15.0) between these same two points. Thus, the standard scores divide up this interval more finely than the grade equivalent scores do. Similarly, for the Total Mathematics Subtest, there are 71 (=95.0 - 24.0) standard score units in the interval between 1.1 and 6.8 on the grade equivalent scale.

Finally, as a purely practical matter, the standard scores are the only form of derived scores available for the Primer Level Test. No grade equivalents have been defined for this test by the publisher.

There is one disadvantage to the standard scores, however. Standing alone, they do not provide the reader with an intuitive appreciation of the magnitude of the gain made by the students. From the point of view of the researcher this is an advantage, because it focuses attention on the relevant comparison between experimental and control, rather than on the less relevant comparison between these students and the national norm. However, it does nothing to assist educators in determining the practical importance of the findings.

It would be desirable if some benchmark were available to indicate the amount of growth to be expected for the test publisher's norm group.

However, this cannot be done with precision. Because there is little available data on "pre" and "post" standard scores on Metropolitan Achievement Tests on large samples of populations such as the one involved in this project, no direct comparisons with a familiar reference group can be made. Also, because the growth pattern in the standard scores is nonlinear (being a negatively accelerated curve similar to those published by Science Research Associates (1967)), there is no direct conversion possible between average growth and the publisher's norms for individuals. In an effort to develop conversion formulas so that a familiar benchmark could be used, several different calculating procedures were created; they each yield somewhat different results. However, as an approximation, it can be stated that for the publisher's norm group, a school average gain of approximately 9 to 11 standard score units would be expected in the time period covered by the project.

With these benchmarks in mind, one can use the standard scores for the analysis, and yet refer to a familiar reference point for determining the educational significance of differences which are found. In the following pages, reference will be made to this benchmark whenever the occasion warrants.

In addition, to further specify the practical magnitude of treatment impact in educational terms, we have calculated the percentage by which the observed gain in standard scores for the experimental group exceeds the observed gain for the control group. This is determined by dividing the difference between the experimental group gain and the control group gain by the control group gain. For example, if the experimental group gain were 7.5, and the control group gain were 5.0, then the calculation would be:

$$\frac{7.5 - 5.0}{5.0} = \frac{2.5}{5.0} = 50\%$$

Although the exact limits are somewhat arbitrary, for purposes of interpreting the educational significance of the marginal gains, the following categories were established:

- 0 to 14.9% net gain is negligible.
- 15 to 29.9% net gain is small.
- 30 to 49.9% net gain is medium.
- 50% and more net gain is large.

Thus, in the example just presented, the EXP school net gain of 50% would have a "large" educational significance. This line of reasoning seems consistent also with the fact that the rate of gain for students in control schools of this sample is approximately 65-70% of the rate of gain for the publisher's norm group. Thus, a large effect in our terms would be one which brought the gain rate of the target population to be nearly equal to that of the publisher's norm group, and this would be almost exactly a 50% increase in gain rate.

A second matter that has to be decided is the exact nature of the comparison statistics to be used. For the present report, the primary mode of comparison adopted was the student t-test. This test is carried out by pooling the data from all students (or teachers or parents) who are properly included in one side of the comparison, finding their average and standard deviation, and then doing the same for all data properly included in the other side of the comparison. Then, these two means are compared, with their standard deviations being assumed equal in the population, and a t-test is calculated. The resulting t-test is then presented along with the means, standard deviations, and number of cases, in a summary table.

A technical feature inherent in this procedure is that in some instances it may not be the most efficient test possible. That is, it may leave as part of the error variance (via the standard deviations used in the t-test) some components which could have been extracted. Although this fact is recognized, it is not regarded as a weakness. The reasoning is as follows. First, for most of the comparisons of direct interest, and particularly for those in which it might be possible to extract variance due to other components (*e.g.*, grade, site), the number of cases in the

comparison is so large that there is no real likelihood of missing an educationally sizable effect because of the limitations of the significance test. On the contrary, several of the results turn out to be statistically significant, but so small in practical terms that there is no point in pursuing them any further. Thus, the emphasis in this project is not on extreme precision of statistical testing, but for good estimates of the <u>size</u> of educational effects and for understandable procedures.

The t-test used here is a statistical technique that is usually covered in first courses in statistics and so is familiar to most persons who will be reading this report. Its application here is an entirely justifiable one because all of its assumptions are easily met (at least in the analysis of the achievement test data). The distributions expected from these tests are symmetrical and approximately normal, the variances are almost certainly equal in the population, and the samples being compared are always statistically independent and not too different in a number of cases.

In addition to the direct analysis of differential gains, a supplementary analysis was also performed on the achievement test data. This supplementary analysis involves again the use of the t-test, but the test was applied separately to the pretest data and to the post-test data. This supplementary analysis is presented in the section following the main analysis.

Organization of Findings

In terms of the broader organization of the hypotheses which was mentioned earlier, all of the material in the main and supplementary analyses of this section deals with achievement test data as a variable and with students as the persons being examined. These analyses are oriented to a set of comparisons developed out of the combined work of OE, TAC, and MAC personnel. They represent a distillation of ideas arising from direct policy interest and those arising as a result of substantive questions about the possibility of "site-by-treatment" interactions. They are summarized in Table III-2 and described on the following page.

TABLE III-2

LIST OF COMPARISONS MADE TO EXAMINE PROJECT RESULTS

Single-Site Comparisons

1. Experimental school versus Control school, Cincinnati, all grades combined.

2. Experimental school versus Control school, Jacksonville, all grades combined.

3. Experimental school versus Control school, Oakland, all grades combined.

4. Experimental school versus Control school, San Antonio, all grades combined.

Model-Based Comparisons

5. Experimental school versus Control school, Teacher-Only Model (Jacksonville and Cincinnati).

6. Experimental school versus Control school, Parent-Teacher Model, (Oakland and San Antonio).

7. Parent-Teacher Experimental schools versus Teacher-Only Experimental schools (Oakland and San Antonio/Jacksonville and Cincinnati).

7a. Parent-Teacher Experimental and Control schools versus Teacher-Only Experimental and Control schools

8. Experimental schools, all cities, versus Control schools, all cities.

Comparisons 1, 2, 3, and 4 in the above table involve pooling data from all the grades at each school at each individual site and comparing the results of the experimental school at a given site with those of the control school at the same site. These four comparisons represent the most direct and focused test of the hypotheses about the impact of the two incentives models that can be made. It should be remembered that the models were applied to the schools as intact units and that in the design of the project there was no policy interest in differential effects of a given model at different grades. If such differential effects do

exist, unless there is a consistent pattern to these effects, (such as greater gains in the lower grade levels across sites), they are irrelevant interactions. To examine these questions, separate tabulations have been run for each grade on the achievement test data. The findings of these grade-by-grade analyses are discussed later in this report. However, these grade-by-grade results are of only secondary interest. The comparisons of what is called the "EXP" (for experimental) and "CON" (for control) entire schools at each site convey the principal results of the project. Thus, the single-site comparisons make up the first four of the summarizing comparisons.

Comparisons 5 and 6 aggregate the data by the type of model used and compare combined EXP against combined CON. Comparison 5 pools the data from the two sites at which the Teacher-Only model was used (Cincinnati and Jacksonville) and compares the results in those two EXP schools with the results in the two corresponding CON schools. Comparison 6 involved a similar comparison of the Parent-Teacher model at two sites (Oakland and San Antonio). These comparisons offer a good example of why statistical elegance is not needed here, and would be distracting if used. The main question, as already indicated, concerned the effect of the treatment models. There are two kinds of treatments, the PT model and the TO model. To assess the overall effect of the TO model, one must examine the results in the two sites in which it was applied. If those results are similar in both sites, then it can be concluded that there was no "site-by-treatment" interaction for the TO model, and the two impacts can be averaged for a more precise estimate of the overall impact of the TO model. If there is an effect at each of the two separate sites, and that effect is statistically significant as well as substantively nontrivial, then it is inevitable that the average of those two effects also will be statistically significant.

On the other hand, if the sites have impacts that differ widely from each other, then an averaging of their effects will not be completely interpretable; hence, it will be less useful. This situation would arise because averaging suppresses an aspect of the phenomenon, namely site-to-site differences in responsiveness, that <u>is</u> of major policy interest.

Under the circumstances that there are site-to-site differences in responsiveness to the incentives, the averaging together of two sites is at best a very crude indicator of a hypothetical average. The limitations of this pooled average under those circumstances must be strongly emphasized. The sites examined are only two in number, and they are neither probabilistically nor even representatively chosen. Thus, generalizations from them to some larger population of potential sites are almost completely judgmental. Nonetheless, since this is a central focus of policy interest, the model based comparisons are provided.

Comparison 7 was carried out to address the question of the differential impact of the PT model in comparison with the TO model. Obviously, the PT model is an extension of the TO model, albeit a potentially expensive one because incentives are offered to parents as well as teachers. Therefore, from a policy point of view, this model should be seriously considered only if it produced substantial marginal benefits beyond those obtained from the TO model. As one way to address this question, Comparison 7 pooled the two EXP schools in the PT incentives model and separately pooled the two EXP schools in the TO model. It then compared the outcomes in the PT-EXP schools with the outcomes in the TO-EXP schools. In other words, the TO-EXP schools in this comparison were used as the "control" group for estimating the marginal impact of the PT incentives model.

This is intuitively a legitimate comparison, because the PT incentives model includes all the features of the TO incentives model, plus others as well. Note, however, that the comparison just described does not eliminate site-to-site variation in learning rate as a source of error variance and possible bias. Stating this differently, the comparison does not answer the question: Is the net impact of the PT incentives model substantially greater than the impact of the TO incentives model, when site-to-site differences in learning rate are taken into account? To deal with that question, a special comparison was performed on the achievement gains analysis (Comparison 7a in Table III-2); it is described later.

The last comparison is an overall one. It suppresses all differences between the different incentives models, and pools across all four sites. This comparison is in many respects vague and imprecise, but it is the best that can be offered to answer the very broad question that policy interest raised: Do the incentives, considered overall, seem to have a sizable impact? In one sense, this last comparison is oversimplified. Yet, it is included for the benefit of those who wish to see a very general summary comparison. Clearly, this overall comparison would have the most interpretability if the PT treatment showed <u>no</u> additional impact over the TO treatment, and if there were no site-by-treatment interactions. As will be seen in later sections, these conditions do not hold for the data at hand, and so the comparison is less meaningful.

Achievement Data Analysis

Discussion

The basic analysis of the achievement test data was carried out by comparing the observed gains made by students in the experimental school (or schools) with the observed gains made by students in the corresponding control school (or schools). These comparisons are the most direct and straightforward way to examine the impact of the incentives models. Moreover, given the characteristics of the design employed in this project (*i.e.*, the use of intact, paired schools) these direct comparisons are methodologically the most appropriate technique available.

Over the last decade, there has been much debate in the statistical literature of the behavioral sciences as to the appropriate procedure for dealing with change and growth data (Lord, 1956, 1958; Harris, 1963; Bereiter, 1963; Lord, 1967; Coleman, 1968; Werts & Linn, 1970; Cronbach & Furby, 1971). The problem which makes moot any analysis of change data is that for most kinds of data, the amount of change observed is dependent on (and predictable from) the initial level. As Coleman's (1968) discussion indicates, the substantive and methodological sources of that dependence often are quite complex. Thus, the analytical question becomes: What is the most appropriate way to adjust for the impact of differences in initial level, so as to be able to ascertain the impact of other influences on the amount of change? The untangling of all these components of the dependence of growth on initial level, particularly when errors of measurement as well as substantive forces must be taken into account, is difficult. Various lines of reasoning, and various technical approaches, have been suggested and are being explored, but at present it must be said that this whole subject area is one that is still in a state of development by statisticians and methodologists.

In the Incentives Project, however, the issues as to what adjustment is appropriate are not very serious in practical terms. The design of the study, and particularly the choice of schools within each site which were as similar as possible with respect to average achievement level,

make any adjustment largely a matter of refinement. The general direction of differences is not likely to be changed by the small adjustments which might be made.

Although there are some differences between schools within a site as far as initial score is concerned, these differences are largely due to random fluctuations, since the schools have been selected to be similar. These random fluctuations will tend to produce effects which make the treatment impact appear somewhat larger than it is. However, this is a known direction of bias, and so can be taken into consideration when the results are interpreted. For these reasons, it seemed that the simpler analysis without any adjustment of scores--namely, a comparison of observed gains--was preferable to some adjustment that would introduce additional uncertainty. Thus, the direct comparison of observed gains has been used for the main analysis.

However, because some distorting factors may be present, additional analysis techniques with the achievement test data have been employed in some further analysis. These other techniques isolate and adjust for certain of the distorting influences that may be present. The results obtained using these supplementary analysis techniques are reported in the next section of this chapter. In general, they do not provide any grounds for substantially changing the conclusions drawn on the basis of the direct analysis of gains in achievement test score, although they do reduce the clarity of the results and the confidence that can be placed in them. In other words, the best guess as to what happened is the same, but its chance of being right shrinks somewhat.

The results of the direct comparison of achievement test gains are summarized in Tables III-3 and III-4. Table III-3 summarizes the Total Reading results and Table III-4 summarizes the Total Mathematics results. The results shown in Tables III-3 and III-4 were computed using all students for whom Total Reading scores were available on <u>both</u> the fall (pre) and spring (post) test administrations. Each table shows the average gain in score, the standard deviation of the gains, and the number of cases involved for each of the comparisons previously discussed.

TABLE III-3

TOTAL READING GAIN

COMPARISONS	GAINS						STATISTICS		
	EXPERIMENTAL GROUP			CONTROL GROUP			t-val.	r_{pb}	% Net Impact E/C
	$\overline{X}$ Gain	S.D.	N	$\overline{X}$ Gain	S.D.	N			
1. E/C, Cin (TO)	5.71	5.29	409	6.39	5.88	362	-1.69	-.061	-11%
2. E/C, Jax (TO)	6.90	6.05	738	6.44	5.59	693	+1.49	+.040	+ 7%
3. E/C, Oak (PT)	7.14	6.11	361	5.04	5.84	487	+5.09	+.172	+42%
4. E/C, San (PT)	8.07	7.15	333	7.65	6.66	288	+0.76	+.030	+ 6%
5. E/C, TO	6.48	5.81	1147	6.42	5.69	1055	+0.22	+.000	+ 1%
6. E/C, PT	7.59	6.64	694	6.01	6.28	775	+4.69	+.122	+26%
7. PT-E/TO-E	7.59	6.64	694	6.48	5.81	1147	+3.77	+.088	+17%
7a. PT/TO	6.89	6.11	1749	6.29	6.01	1922	+2.99	+.049	+10%
8. E/C, All	6.90	6.16	1841	6.25	5.95	1830	+3.25	+.054	+10%

TABLE III-4

TOTAL MATHEMATICS GAIN

COMPARISONS	GAINS						STATISTICS		
	EXPERIMENTAL GROUP			CONTROL GROUP			t-val.	r_{pb}	% Net Impact E/C
	$\overline{X}$ Gain	S.D.	N	$\overline{X}$ Gain	S.D.	N			
1. E/C, Cin (TO)	8.46	8.03	410	9.57	8.03	358	-1.92	-.069	-12%
2. E/C, Jax (TO)	9.90	7.64	734	8.37	7.37	679	+3.83	+.101	+18%
3. E/C, Oak (PT)	9.95	7.68	357	6.73	7.82	488	+5.96	+.201	+48%
4. E/C, San (PT)	12.72	9.96	334	8.84	7.98	287	+5.30	+.208	+44%
5. E/C, TO	9.38	7.81	1144	8.79	7.62	1037	+1.81	+.039	+ 7%
6. E/C, PT	11.29	8.96	691	7.51	7.94	775	+8.56	+.218	+50%
7. PT-E/TO-E	11.29	8.96	691	9.38	7.81	1144	+4.79	+.111	+20%
7a. PT/TO	9.79	8.27	1728	8.63	7.91	1919	+4.32	+.071	+13%
8. E/C, All	10.10	8.31	1835	8.24	7.78	1812	+6.98	+.115	+23%

Each row of the table presents the results for one comparison pair of groups. Each row also gives the value of student's t and the point biserial correlation coefficient (r_{pb}) between treatment and gain. These point biserial coefficients seem small because they are based on students in all grades, so that any variation between grades is not taken out of consideration. For all the rows of the table, the group whose results are shown on the left is the experimental group (EXP) and the group on the right is the control group (CON). That is, the left side shows results for students in the school (or schools) receiving the incentive model (the presumably more powerful of the treatments).

Thus, the left side of the third row of Table III-3 indicates that among all the children in the experimental school in Oakland, the average gain in Total Reading score from fall to spring was 7.14 points. Similarly, the right side of the third row indicates that for all the children in the control school in Oakland, the average gain in Total Reading score was 5.04 points.

A formal test of significance logic has not been used, because the real policy question involves the magnitude of the differential gains rather than their mere existence. However, the existence of nonaccidental differences in the gains is a necessary condition that must be met before it can be concluded that the treatment has an effect. For the purposes of an analysis of this sort, it is not misleading to use a t-value of approximately 2.00 as a criterion of statistical significance.

As a caution, it should be noted that the existence of differences too large to attribute to sampling fluctuation does not, in itself, guarantee that these differences were caused by the treatment. The differences which are observed in any site are the net result of <u>all</u> the various systematic differences that exist between the two schools. The careful matching of schools in the design of this project has eliminated many of these systematic differences that might otherwise be present. However, it has not eliminated <u>all</u> such differences, and so there is no way to be sure that the outcomes are due to the treatment. Since only a small number of schools were involved, and since evidence in this data

indicates that there are substantial differences among these schools, there can be no appeal to an argument that other differences will be averaged out over the many cases involved. In other words, the existence of a sizable difference in the predicted direction (such as that observed for Oakland in Table III-3) is an encouraging sign, but cannot be taken as conclusive evidence of a treatment impact. Conversely, the absence of such differences in outcomes may likewise be due to the operation of other differences besides those of treatment and control, and so negative conclusions also must be tentative.

Table III-3 presents the results for achievement gains in Total Reading score. This is referred to throughout the report as "Total Reading" or "Reading". Of the nine comparisons in this table, eight show differential gains in the predicted direction (of course, several of these comparisons are statistically dependent on the others). When an unweighted analysis was done, to eliminate any distortion due to differences between schools in distribution of students across grades, seven of the results remain in the predicted direction. The exception was the E/C, TO comparison, row 5, which changed from +.22 to -.50. However, several of the differential gains are quite small, and only five of the Total Reading t-values are greater than 2.00. Thus, the general though tentative conclusion is that the offer of incentives did not have a uniformly large effect on Total Reading gains.

Single-Site Comparisons--Reading Gains

In Cincinnati, the comparison of gains in Reading shows that students in the EXP school on the average gained less than the Cincinnati students in the CON school, but the difference in gain is negligible. Here, we must tentatively conclude that the incentives model had no beneficial effect on Reading achievement.

The second row of Table III-3 shows the comparison for Jacksonville. As can be seen from the table, the differences in gains are negligible although in the direction predicted. In this city as well, we must tentatively conclude that there is no beneficial effect of the incentives model on Reading achievement.

Students in the EXP school in Oakland gained more, on the average, than those in the CON school. Students in the EXP school, on the average, gained 7.14 points, and students in the CON school gained only 5.04 points. This is a difference in gain of 2.10 points, or 42% of the control school gain. A 42% net gain is at the upper end of the "medium" gain category presented on page III-7. Thus, the Total Reading impact in Oakland is clearly too large to be neglected, and it is nearly as large as might be expected if the incentives treatment were as beneficial as deserved.

In San Antonio, the difference in gains is quite small, and the t-value is less than 1.0. The direction of difference is, however, as predicted by the hypotheses. In San Antonio, there is no evidence of a beneficial impact of the incentives model on gains in Total Reading achievement.

Note that, at this point, several complications begin to emerge from the pattern of results. These complications create difficulties over and above the other difficulties already mentioned. Generally, the incentives program seems to have had an erratic impact on Reading achievement. These variations in the impact complicate the drawing of any concrete policy inferences. For instance, there is clearly a "site" effect on Reading gains. The gains in San Antonio, for <u>both</u> the EXP and CON schools, are greater than the gains in Oakland. A number of possible explanations for this can be offered. These differences in achievement gain may be due to differences in the school characteristics, or differences in the cities, or differences in the region. Without considerable additional data, there is little that can be done to account for these differences, although the process evaluation evidence (Education Turnkey Systems, 1972b, p. 65) suggests that the administrative "climate" in San Antonio was more favorable than that in Oakland, and this may be at least a partial explanation.

Clearly, the variation in gain between untreated schools in different cities is approximately as large as the impact of the treatment, and so the treatment cannot be universally recommended, at least not without much additional understanding of the conditions that cause the natural variations in gain rate.

The site-to-site differences in gain rate have other consequences as well. Because of the small number of sites used, and the fact that not all three possible treatment levels were used in any site, (*i.e.*, the PT and TO treatment levels were never used in a single site together) there are problems introduced in drawing inferences about the relative effectiveness of the two incentives models. In particular, differences between the PT and the TO treatments are not easily distinguishable from differences between sites in receptivity to the treatment. That is, the contrast between PT and TO treatments is confounded with the treatment-site interaction.

The design used in the project controlled for differences between sites in the composition of the student body, as far as initial score level is concerned, but it did not control for differences in the current learning rate or receptivity among the four sites. If the sites where the PT treatment was used happened to be operating with a higher general learning rate or were more receptive than those at which the TO treatment was employed, then in Comparison 7 there is likely to be an apparent positive effect of PT which is in reality due to differences in sites.

Also, within the pair of sites that received the same incentives treatment there may also be differences between sites in their responsiveness to the treatment. This seems to be the case for Oakland and San Antonio with the Parent-Teacher incentive treatment. It appears that Reading scores in Oakland respond to the treatment rather substantially, while San Antonio seems essentially unaffected. Because of this treatment-site interaction, any broad conclusions as to the effectiveness of a particular incentives model become quite tenuous. That is, the combining of Oakland and San Antonio experimental schools into a single entity called the Parent-Teacher model population would lead one to draw conclusions only about what must be thought of as an imaginary average effect of the Parent-Teacher model.

There are two difficulties with the averaging concept. First, the average may be meaningful for policy making at the broadest Federal levels, but it raises at least as many questions as it answers, because there is

no way to predict what the impact of the Parent-Teacher model will be for a given particular school. Second, the application of an average effect idea is weak here, because the set of units over which the average is calculated (two schools, one in Oakland and one in San Antonio) is small and unrepresentative of any politically meaningful group.

Perhaps the most important inference that can be drawn from this discussion is that there are some very important factors--presumably ones operating at the school level, but perhaps also others being determined by neighborhood or school system characteristics--that have been completely neglected in the design. This omission is not the result of negligence, but rather of ignorance. The field of urban education simply does not offer firm knowledge about the particular variables, or combinations of variables, affecting the success of a school. Moreover, much of the available evidence indicates that many of the seemingly important factors are nonmanipulable by any foreseeable kind of policy change. In short, the results here serve best perhaps as an object lesson in the primitive state of knowledge about the factors contributing to school effectiveness. Because such knowledge is so limited, the other factors which affect either the general learning rate or the responsiveness to an incentives treatment have neither been held constant, nor have they been measured so that they can be adjusted for in the calculations. Yet, the data indicate that such influences are substantial in size.

Also, the possibility was examined that the differential effects may arise out of a combination of the score format used (in which the typical gain is larger at the lower grades than at the upper grades) and the distribution of students across grades in the different schools. To check on this possibility, a separate analysis of the achievement data was executed, using an unweighted means analysis (Winer, 1962, p. 222-224). This analysis effectively makes the distribution of students across grades the same for every school, and so makes it impossible that differences in such distributions could be contributing to the effects observed. These results are not included here, for lack of space, but they are very similar to the main analysis, and provide no reason to change any of the conclusions drawn.

Model-Based Comparisons--Reading Gain

The qualifications mentioned above are serious ones. From a methodologically conservative point of view, they would be sufficient reason not to perform any further comparisons. Yet, these other model-based comparisons are not difficult to carry out, and may, if interpreted with the necessary caution, be of some use to policy makers and research planners. For that reason, they have been calculated and are presented below.

The first of the model-based comparisons was made by pooling the two sites which experienced the TO model. For these two sites, the direction of effect differed, but the magnitude of the effect was negligible in both sites. When the two sites were pooled, these two opposite directions tended to cancel one another, and the resulting average gain in the EXP schools, 6.48, was nearly identical to that in the CON schools, 6.42. Thus, on balance, it would seem that the TO model has no discernible impact on Reading achievement in this data.

The next model-based comparison was constructed by pooling the two sites which experienced the PT treatment, and comparing the EXP schools with the CON schools. Here, the impact was in the predicted direction in both sites, although it was much larger in one site (Oakland) than the other (San Antonio). When the two sites were pooled, the average gain in the EXP schools was 7.59, and the average gain in the CON schools was 6.01. This would seem to suggest, on balance, a non-negligible, but small, impact in the sites exposed to the PT treatment. The t-statistic of +4.69 is consistent with this interpretation, and indicates that the difference almost certainly is not due to chance statistical fluctuations. Of course, as mentioned earlier, such differences may be partly due to differences between schools other than the treatment. In terms of impact, the net gain was small, the PT-EXP schools gaining 26% more than the PT-CON schools.

The next comparison is a questionable one. It was aimed at contrasting the impact of the PT model with the impact of the TO model and made use only of the EXP students. It compared, on the one hand, students from EXP schools in Oakland and San Antonio with, on the other hand, students

from EXP schools in Cincinnati and Jacksonville. Thus, differences which are found may be due to differences in the two incentives models, to differences in the "main" effect of sites, or to an interaction between the two factors. In short, it is a muddy comparison. However, it does show an apparent effect. That is, the average gain among the EXP students in the two sites which were exposed to the PT model is 7.59, and the average gain among students in the EXP schools in the two sites which were exposed to the TO model is smaller, only 6.48. Although the t-statistic was 3.77, the net impact was only 17%, which is on the borderline between "small" and "negligible".

A more complicated comparison of the PT versus TO models can be constructed which eliminates the confounding between the site-to-site differences in growth and the net impact of the PT treatment. This comparison (7a in Table III-2) includes taking the difference in gain between the EXP and the CON for the PT sites, and comparing it with the difference in gain between the EXP and the CON for the TO sites. When this comparison is carried out, the difference in gain is no longer 1.11 (7.59 - 6.48), but instead is only 0.60 (6.89 - 6.29). This amount is still large enough to have a t-value of 2.98. However, it is too small to be of any practical significance.

The last comparison presented in Table III-3 is an overall comparison of the gains in Reading score in all four experimental schools compared with the gains in all four control schools. It does not address any precise hypothesis about the processes that are operating, but instead gives some general sense of the magnitude of the average impact that was achieved. As can be seen, the difference in gains between the EXP and the CON schools is not large (0.65).

Although the t-statistic was 3.25, this is an educationally negligible difference, for the net gain was only 10%.

Single-Site Comparisons--Mathematics Gains

Table III-4 presents the results for achievement gains in Total Mathematics score. This is referred to throughout the report as "Total Math" or "Math". It should be noted at this point that the Math test score gains, in almost all instances, are larger than the Reading test gains. This is simply a function of the content of the test, and does not indicate anything at all about the quality of instruction or relative performance in the two subjects. There is no way to compare these scores across subject directly. Again, the results in each of the four sites are considered individually.

In Cincinnati, the comparison of the gains in Math indicates the students in the EXP school on the average do not gain quite as much as the students in the CON school, but the difference in gain is negligible (-12%). The amount and direction of the Total Math results are similar to the results obtained in Cincinnati for the Reading tests.

For Jacksonville, the average gain among the students in the EXP school is 9.90 points, and the gain among students in the CON school is 8.37 points, yielding a difference of 1.53 points. This impact, although statistically large enough to be real (t-value of +3.83), produces an impact of only 18% which is just above the cutoff point between the "small" and "negligible" gain categories.

The EXP school in Oakland gained noticeably more, on the average, than the CON school in Oakland. The gains were, respectively, 9.95 and 6.73. The difference is 3.22 points and is clearly statistically significant (t-value is 5.96). The net impact on the EXP school is 48% and is on the border between "medium" and "large". Thus, the Total Math results in Oakland are similar to the Total Reading results for Oakland.

In San Antonio the EXP students gain, on the average, 12.72 points, and the CON students gain, only 8.84. This is a difference of 3.88 points and has a t-value of +5.30. The net gain is 44%, indicating a medium-to-large impact of the treatment (PT) in San Antonio. This is a clear instance of a substantial percentage impact of the treatment on the Total Math results, but not on the Reading results, in San Antonio.

Model-Based Comparisons--Mathematics Gains

Although Comparisons 5 through 8 in the table address hypotheses originally formulated in the design of the project, they must be interpreted very carefully in light of the patterns observed in the Total Math data. Here as in Reading, the gains in San Antonio, for both the EXP and the CON schools, are greater than the gains in Oakland. Also, there are again some noticeable differences in response to the treatment between the two sites which are given the same treatment. For example, Cincinnati responds negligibly and in a negative direction to the TO treatment, but Jacksonville responds somewhat positively (though not enough to be important educationally).

As before, these patterns seriously restrict the generalizations that can be made. Yet, the model-based comparisons have been reported so that policy makers at all levels may have an understanding of the complexity inherent in interpreting the results. Much caution should be used in generalizing to policy interpretations from this set of findings.

The comparison for the two sites which experienced the TO model indicates a gain among the EXP group of 9.38 and a gain among the CON group of 8.78, with the difference between these being 0.60. Although this is statistically large enough to be considered, it is negligible educationally. Thus, the data indicate no impact of the TO treatment, on balance, keeping in mind the treatment-site interaction here.

The next model-based comparison involves the two sites which were given the PT treatment. Here, there is a real impact. The gain among the EXP school was 11.29 points, and that among the CON schools was 7.51 points. The difference is 3.78 points. The t-value is +8.56; this is clearly large enough to be unlikely to be due to fluctuations. Furthermore, the net impact is 50% which is large enough to justify further exploration. It should be noted that the Total Reading results for the PT comparison had a t-value of +4.69 and a net impact of 26%. Thus, the two findings are similar in direction; together they indicate a possible pattern of favorable PT model results.

The comparison labeled "PT-E/TO-E" is one that must be viewed with some caution. It reveals that the difference in gain between the EXP students in the PT sites (11.29) and the gain of the EXP students of the TO sites (9.38) is 1.89 = 20%. This is large enough to have a t-value of +4.79, but the net impact is only 20%. This comparison is possibly distorted by differences between sites in the prevailing growth rates. To cope with this, the same comparison that was used with the Reading results (see pages III-22,23) was employed. This comparison yields a difference of 1.16 between the gains of the PT and those of the TO, after adjustment for site-to-site differences in growth rate. Thus, the net impact is reduced to 13%. This impact is just large enough to be considered non-negligible. It clearly is above any necessary cut point in the t-statistic, having a t-value of 4.32.

The last comparison is, as with the Total Reading results in Table III-3, an overall comparison of all four experimental schools with all four control schools. It indicates that the average gain among the EXP schools is 10.10 points and the average gain among the CON schools is 8.24 points. Thus, the difference in gains is 1.86 points. This again has a high t-value, +6.98, but the educational relevance of this impact is small, 23%.

Summary

To summarize the results of this analysis of achievement gains, it seems reasonable to say that some of the hypothesized differences do appear in the data. For achievement in Reading, the incentives models appear to have had an impact worthy of note only in Oakland. While the comparison for the sites receiving the Parent-Teacher incentives model indicates an impact on Reading that is large enough to consider further, it must be remembered that the effect of the PT model was not similar in the two sites where it was employed. The model appears to have worked fairly well in Oakland, but not at all in San Antonio.

For Mathematics, the picture is somewhat clearer. In both Oakland and San Antonio, the incentives model (PT) seems to have had a sizable

impact on achievement gain. However, the TO model sites do not show any detectable effect of the incentives model they experienced. One possible explanation for this difference in the effects of the PT model on Reading and on Mathematics might be that parents find it easier to understand the difficulties a child is having with mathematics, and to provide him with useful help. This is, of course, only a possibility, but perhaps is worthy of further investigation.

Any general conclusions from this project must be drawn very cautiously, if at all, since the data indicates considerable site-to-site differences in achievement level, gain rates, and responsiveness to the models. This perhaps is as important as any other result of the project. Furthermore, supplementary analyses presented in the next section discuss some additional sources of "noise" in the data, which tend to reduce the definiteness of any conclusions that can be drawn.

Further Analyses of Achievement Test Data

Introduction

The results presented in the preceding section provide the most direct and clear summary possible, of the impact of the incentives models on academic achievement. While the mode of analysis is a simple one, it is as satisfactory as any which might be used, given the characteristics of the project design and the data. However, there are some additional points that should be examined and some methodological explorations which might support the previous results or reveal needed qualifications. Thus, some additional analyses were carried out and are described in this section.

The use of observed gains in MAT standard score rather than some form of adjusted gains such as covariance analysis was adopted after a review of the literature and a consideration of the issues involved. First, the use of observed gains provides no loss in the reliability, as compared with any form of adjusted gains, because the reliability of the adjusted gains also depends on the reliability of the data at each time point. More importantly, the analysis here is of <u>average</u> gains, the measurement reliability of which is considerably higher than that of individual gain scores. Also, because the interest is on differences in average gain only, the whole issue of the dependence of individual gain on the initial score of the individual is not the problem. Rather, the question is whether the <u>average</u> gain is dependent on the <u>average</u> initial score. If that is the case, the next question is whether adjustments should be made for that dependence. If the answer to that question is in the affirmative, then there still remains the problem of calculating an appropriate coefficient to carry out the adjustment.

Ideally, it would be preferable to adjust appropriately for any dependence of average gain on average initial score. However, with only four comparable units, the data are not sufficient to obtain any precise estimate of the form of dependence. Hence, a judgmental argument is used instead. The conclusion from that argument is that the dependence

of average gain on average initial score, between schools within site, will be negligible. This conclusion is derived from the following two arguments. First, the schools being compared begin at nearly the same point on the general growth curve of Reading achievement or Math achievement and are otherwise quite similar. Consequently, their natural growth rates can be expected to be very similar. In other words, there is no conspicuous relation between initial score and growth rate arising from this source, at the school level of comparison. The "fan spread" will be slight. Second, there is no particular reason to expect a sizable regression of the two <u>school means</u> within a site to a <u>site mean</u>. That is, the small achievement differences between schools within site which do exist are probably stable over time. This conclusion is based only on the limited evidence culled from the districts' past achievement test results, but it is nonetheless the best interpretation available in the absence of richer and more appropriate data.

So, then, the decision not to adjust is based on evidence which is not conclusive. However, it is the best evidence available. An additional reason for not adjusting is that there is no satisfactory way of deciding what value the adjustment coefficient should have. If a coefficient were chosen on the data available, and adjustment were made according to it, there would be no way of knowing the direction of the error in the resulting adjusted score.

Before going on to the specific analysis, one further point should be made about the whole question of comparing gains in academic achievement. There are two particular ways in which growth can depend upon initial score--one is "fan spread", and the other is regression toward the mean. These factors obviously complicate any attempt to control for the dependence of growth on initial score. There is an additional complication, however. The two sources of dependence tend to work in opposite directions. The use of linked interval scores such as the MAT standard scores employed in this analysis reduces this problem somewhat for school level analysis by counteracting the positive dependence of growth on initial score across age levels. In other words, through use of the standard

scores, there can be more confidence that the direction of any small net bias in school average gains is to make the initially low scoring group show a greater gain. So, while there is still some ambiguity present in the results, the combination of influences almost certainly will be such as to make the final results (if anything) slightly optimistic as far as estimates of treatment impact are concerned.

To pursue this and other matters in more detail, some methodological investigations have been carried out. Several topics were considered in these further analyses:

1. The fan-spread phenomenon and ways to deal with it
2. The regression toward the mean phenomenon and its relevance in this analysis
3. Patterns of achievement gains within each grade
4. Specific patterns in the results that can be identified when starting and ending scores as well as gains are studied.

Fan-Spread Analysis

As discussed previously, one problem that has been noted in the use of achievement test data is that the amount of gain which occurs in the scores of target populations is to some degree proportional to the initial score level of those populations, even in the absence of any treatment at all. Thus, if a study is carried out in which the control group has initial scores higher than those of the experimental group, the simple comparison of gains made by the experimental group with gains made by control group would tend to be biased against finding a positive impact from the treatment. This point was made with slightly different emphasis by Campbell Erlebacher (1970) in their review of the Westinghouse-Ohio University evaluation of Head Start. It is also discussed in some detail by Campbell in a paper published in the 1970 Proceedings of the Educational Testing Service Invitational Conference on Testing Problems.

The basic approach suggested by Campbell in the Educational Testing Service paper is to note the fact that if growth rates are, in the absence

of any treatment, proportional to initial score, then the variance of a set of scores at Time 2 will be greater than the variance of the same scores at Time 1, by the same proportion. In other words, Campbell notes that this difference in variances has been observed empirically and he makes the connection between it and differential growth rates <u>within</u> the groups as well as <u>between</u> the groups. The question then is: How does one adjust for these irrelevant differences, so as to be able to obtain an appropriate measure of the treatment's effect?

The idea suggested by Campbell is to compare not the observed gains, which may be affected by differences in the metric (*i.e.*, different standard deviations implying different natural growth rates), but rather to compare the difference between the experimental and the control group at Time 1 (before the initiation of the treatment) with the difference between the experimental and the control group at Time 2 (after the impact of the treatment). The important point is that this comparison is in a metric which is unaffected by differences in growth rate between the experimental and control groups (assuming that such differences are proportional to the standard deviations).

One metric which accomplishes this is that of the t-test. That is, the differences between the experimental and the control groups at Time 1, measured in units of the t-statistic, should be the same as the difference between the experimental and the control groups at Time 2, also measured in units of the t-statistic, provided that there is <u>no</u> real treatment impact, and regardless of whether the variances at Time 2 are larger than those at Time 1. Unfortunately, Campbell does not provide in his articles the exact form of a statistical test which might be used to decide whether or not the two t-values are equal. However, this is not too important, given the data, because the objective is only to get some approximate sense of the size of the differences in the t-value. Since the values being compared are t-values, it seems reasonable to regard differences in t-value of 1.0 to 2.0 as indicating a small change, those smaller than that as negligible, and those larger as being noteworthy.

If this comparison is to be accomplished in an effort to minimize any bias due to fan spread, it is necessary to show the pre and post-test school average results for Reading and Math achievement. Tables III-5 through III-8 do this for Reading Pre, Reading Post, Math Pre, and Math Post, respectively.

The first point that can be noted upon inspection of these tables, which show the standard deviations, case bases, and t-values as well as the school means, is that the standard deviations for the post-tests are all about the same size as those of the corresponding pretest, and usually are slightly smaller. This indicates that the results of the original gains comparisons are not likely to be affected by differences in growth rates between schools.

To be thorough, however, the differences in the t-values from pre to post for both Reading and Math were calculated and each difference was examined. The results of these separate comparisons, as well as summary results from the gains analysis, are presented in Table III-9. They provide no reason to modify any of the conclusions drawn in the main analysis.

Regression Effect

A second possible source of distortion in an analysis of this general type is the elusive phenomenon known as regression toward the mean. Basically, this phenomenon can be described by saying that if a set of scores is examined, and the units with low scores are compared to those with high scores, then there will be a tendency for the high scores to move downward toward the mean of the entire set, and for the low scores to move upward toward the mean of the entire set. If there is a general growth process occurring, this regression phenomenon will manifest itself as a <u>greater</u> growth among those units which were initially low in score.

(Text resumes on page III-38)

TABLE III-5

READING ACHIEVEMENT, MAT STANDARD SCORES, PRETEST

COMPARISONS	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
E/C, CIN	49.63	15.52	409	47.19	15.77	362	+2.16	+.077
E/C, JAX	39.41	15.29	738	46.31	15.44	693	-8.49	-.219
E/C, OAK	47.58	17.23	361	51.45	16.45	487	-3.32	-.113
E/C, SAN	44.34	16.47	333	47.62	18.55	288	-2.33	-.093
E/C, TO ONLY	43.06	16.12	1147	46.61	15.55	1055	-5.26	-.111
E/C, PT ONLY	46.03	16.93	694	50.03	17.35	775	-4.46	-.116
PT-E/TO-E	46.03	16.93	694	43.06	16.12	1147	+3.76	+.087
E/C, ALL	44.18	16.49	1841	48.06	16.42	1830	-7.14	-.117

TABLE III-6

READING ACHIEVEMENT, MAT STANDARD SCORES, POST-TEST

COMPARISONS	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
E/C, CIN	55.34	15.08	409	53.57	14.06	362	+1.67	+.060
E/C, JAX	46.32	15.46	738	52.75	15.34	693	-7.90	-.204
E/C, OAK	54.73	15.85	361	56.49	15.47	487	-1.62	-.056
E/C, SAN	52.41	15.95	333	55.27	16.91	288	-2.16	-.087
E/C, TO ONLY	49.53	15.91	1147	53.03	14.91	1055	-5.31	-.113
E/C, PT ONLY	53.62	15.93	694	56.03	16.02	775	-2.90	-.075
PT-E/TO-E	53.62	15.93	694	49.53	15.91	1147	+5.33	+.123
E/C, ALL	51.07	16.04	1841	54.30	15.45	1830	-6.22	-.102

TABLE III-7

MATH ACHIEVEMENT, MAT STANDARD SCORES, PRETEST

COMPARISONS	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
E/C, CIN	54.49	19.53	410	52.10	20.86	358	+1.64	+.059
E/C, JAX	43.61	20.20	734	51.26	20.98	679	-6.98	-.183
E/C, OAK	50.01	20.94	357	55.27	20.80	488	-3.62	-.124
E/C, SAN	48.98	20.96	334	52.74	21.42	287	-2.21	-.088
E/C, TO ONLY	47.51	20.63	1144	51.55	20.93	1037	-4.53	-.096
E/C, PT ONLY	49.51	20.94	691	54.33	21.06	775	-4.39	-.114
PT-E/TO-E	49.51	20.94	691	47.51	20.63	1144	+2.00	+.047
E/C, ALL	48.27	20.76	1835	52.74	21.03	1812	-6.47	-.106

TABLE III-8

MATH ACHIEVEMENT, MAT STANDARD SCORES, POST-TEST

COMPARISONS	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\bar{X}$	S.D.	N	$\bar{X}$	S.D.	N	t-val.	r_{pb}
E/C, CIN	62.95	17.36	410	61.66	18.66	358	+0.99	+.036
E/C, JAX	53.51	19.21	734	59.63	18.86	679	-6.04	-.159
E/C, OAK	59.96	18.30	357	61.99	18.33	488	-1.60	-.055
E/C, SAN	61.70	17.69	334	61.59	19.13	287	+0.08	+.000
E/C, TO ONLY	56.90	19.11	1144	60.34	18.81	1037	-4.23	-.090
E/C, PT ONLY	60.80	18.02	691	61.84	18.62	775	-1.09	-.028
PT-E/TO-E	60.80	18.02	691	56.90	19.11	1144	+4.33	+.100
E/C, ALL	58.37	18.79	1835	60.98	18.74	1812	-4.21	-.069

TABLE III-9

READING

	Original t For Gains	Original Net Impact	Diff Of t's
CIN	-1.69	-0.11	-0.49
JAX	+1.49	+0.07	+0.59
OAK	+5.09	+0.42	+1.70
SAN	+0.76	+0.06	+0.17
TO	+0.22	+0.01	+0.05
PT	+4.69	+0.26	+1.56
PT-E/TO-E	+3.77	+0.17	+1.57
E/C, All	+3.25	+0.10	+0.92

MATHEMATICS

	Original t For Gains	Original Net Impact	Diff Of t's
CIN	-1.92	-0.12	-0.65
JAX	+3.83	+0.18	+0.94
OAK	+5.96	+0.48	+2.02
SAN	+5.30	+0.44	+2.29
TO	+1.81	+0.07	+0.30
PT	+8.56	+0.50	+3.30
PT-E/TO-E	+4.79	+0.20	+2.33
E/C, All	+6.98	+0.23	+2.26

The possible relevance of the phenomenon in this experiment is as follows. For most of the sites and most of the grades, the students in the EXP schools had average initial scores somewhat below the average scores of the students in the CON schools. Therefore, to the extent that the regression phenomenon operates between schools, within a site, the students in the EXP schools would show greater average gains, regardless of any treatment. In other words, the greater average gains that have been regarded as being evidence of the impact of the models may in fact be simply the result of a statistical artifact. The problem then becomes one of estimating the size of this regression, if possible.

It is here that the issue becomes complicated. The possible scores of the regression effect include all components of the score which are stochastically unstable over the time interval under consideration. That is, the sources include what is ordinarily classed as measurement error and other components as well. In this instance, the portion of the regression effect that is due to measurement error will be extremely small, because the averages under consideration are based on large numbers; and the measurement error of these averages is trace. However, there are other potential sources of regression effects extant, in the form of temporary influences that operate at the grade or school level. These influences are not well understood, as to source or magnitude, but appear in general to be too large to neglect entirely. Thus, a problem does exist. Although the regression of individual scores toward their school averages will not be involved, the regression of the school averages within a site to a common site average is possible and will be confounded with the impact of the treatment.

The reason that this regression of school averages is confounded with the treatment impact is that the schools were not assigned to the experimental or control condition by a strictly random process, but instead by an administrative decision which took into account past achievement scores, and poverty levels in the schools and attempted generally to apply the treatment in the more deprived school. This decision was in accord with the political realities that attend the allocation of school district

resources. This is an unfortunate limitation of most experiments and demonstrations of this type, and it restricts seriously the kind of conclusions that can be drawn. If there had been several replications of the two treatments within each site, then some empirical estimate of the regression of the school means might have been obtained from the convergence of the means of the schools within each treatment. However, such data are not available.

Thus, one is left with no satisfactory way to make a quantitative estimate of the size of the regression distortion. In general, it is clear that the estimate of the treatment impacts presented in the previous section will need to be lessened, given the possibility of at least a small regression of school averages.

A final consideration in this discussion of regression, while far from rigorous, does seem to have some merit. The implementation of the incentive models was far from optimal. The district grants were delayed by three months and the agreements with the teacher and parents were delayed by as much as five months. This is not a criticism, for the logistical and administrative problems encountered were enormous. The point is that the delay in fully implementing the incentives models almost certainly attenuated the potential effect of those models. Thus, this weakness in implementing the models will tend to compensate for the distortions introduced by the regression effect.

The overall conclusion of this discussion of fan spread and regression, then, must be that the specific estimates of impact magnitude reported in the preceding section cannot be regarded as very precise. They might well be off the mark and are probably somewhat high. On the other hand, there is no feasible way to make a quantitative judgement as to how far they are wrong in any specified case, and so the results presented earlier are the most reasonable of any single estimate. The real point to be drawn from this discussion is that firm inferences about the sizes of effects can be made only if appropriate procedures are followed so that there is tight execution of a sound research design.

Grade Level Comparisons

The next point to be discussed concerns the results for the achievement test within each grade. The incentives models were applied uniformly to all grades within the experimental schools. There is no feature of the design, nor any clear theoretical justification, for expecting the impact of the incentives models to differ according to grade levels. Thus, this comparison is not central to the policy issues inherent in the design of the project. However, it does have some relevance to the Federal interest in early intervention strategies and future field applications of incentives.

For those reasons, an analysis of the achievement test results was carried out within each grade and site. These results are summarized in Table III-10, for Reading, and Table III-11, for Mathematics.

The primary distinction to be made about the grade-by-grade analysis is to be clear as to what interpretation is to be given to various possible outcomes. First, if there is no grade-treatment interaction in a particular site (*i.e.*, the impact of the treatment is uniform across the grades), then that is the end of the matter. Alternatively, it may be that there are differences in responsiveness between the grades, and that such differences have a discernible pattern. For instance, the treatment may have a larger impact at Grades 1 and 2 than at Grades 5 and 6. If this were the case, there would be clear scientific and policy implications. For policy, the implications would be a restricted applicability of the incentives models, and for science there would emerge the question of explaining why the grades differ in responsiveness. A third possibility is that differences will be found between grades in their responsiveness, but that these differences will be more or less random. Such a lack of pattern would presumably reflect the operation of idiosyncratic factors.

If the impacts appear to be more or less random, it is reasonable to group the sites within the incentive models. That is, the set of grades being studied (Grades 1 - 6) is quite important in its own right--in statistical terms it is a "fixed" effect--and there are enough cases in each grade so that one can be at least reasonably confident that the

TABLE III-10

READING

GRADE	CIN	JAX	OAK	SAN
1	-1.86	-1.58	+1.26	-3.69
2	-1.93	-4.40	+4.77	+0.90
3	+0.20	+2.71	+3.12	+1.21
4	+1.27	+1.74	-1.23	+0.90
5	+2.58	-0.06	-0.43	-0.04
6	-1.76	+1.72	+3.78	+2.05

TABLE III-11

MATHEMATICS

GRADE	CIN	JAX	OAK	SAN
1	+0.45	-0.54	+2.74	+2.68
2	+0.53	-1.60	+2.19	-0.24
3	-2.46	+1.86	+3.00	+2.17
4	-0.77	+1.43	+0.66	+5.87
5	+1.96	-0.42	+1.37	+1.24
6	-0.37	+4.79	+4.56	+1.35

averages over all the grades are stable. In short, the emergence of the last pattern--random differences among grades in their responsiveness to the treatments--would leave some scientific questions open, but would not obstruct the drawing of policy recommendations.

An inspection of Tables III-10 and III-11 will indicate that indeed it is this third pattern of random impacts which emerges in the data. That is, the impact of the treatment is not consistent within the different grades at a given site, but these differences show no particular pattern when examined across sites.

For example, the overall result for Reading in Cincinnati is a negligible difference in the direction opposite from the hypothesis. Table III-10 indicates that this arises from a situation in which Grades 1, 2, and 6 show negative results, and Grades 3 and 4 show negligible results in a positive direction, and Grade 5 shows definite positive results. When the corresponding data for Jacksonville are examined, one again finds negative results in Grade 2, and results in a negative direction (though negligible in size) in Grade 1. This slight similarity to Cincinnati disappears, however, when the other grades are examined. In Jacksonville, Grade 6 shows a positive impact, as also does Grades 3 and 4. Grade 5 (which was a positive impact in Cincinnati) is essentially nil in Jacksonville. Further inspection of the table reveals no consistent pattern of impacts in any of the sites in either Reading or Math achievement.

Thus, the grades can legitimately be pooled in order to obtain an overall estimate of the impact at the site, and the scientific point of differential impact can be noted. This differential impact is reminiscent of the site-to-site differences noted in the main analysis, a further reminder that there exists a paucity of knowledge about the factors influencing educational productivity.

Specific Patterns in the Achievement Results

The next topic to be discussed in this section is the specific results found at each of the sites. As an aid to the comprehension of these results (which are presented in Tables III-5 - III-8), the same results are presented graphically in Figures 1 - 16.

Each figure shows one of the comparisons in Tables III-5 - III-8, for the pre and post-tests.

The specific results vary from site to site and comparison to comparison. For example, consider the fact that one of the PT sites (Oakland) shows a sizable impact in Total Reading, but the other PT site (San Antonio) shows no detectable effect. These impact differentials are attributable not to differences in gains made by the experimental schools; rather, they are due to the fact that the gain in the Oakland control school was the lowest of the four schools under consideration, whereas the gain in the San Antonio control school was the highest of the four. This observation does not directly explain why this difference occurred; but perhaps it, like the other results, can shed some light on the vagaries of field research in education. In other words, the significance of the finding for achievement gain in Reading should not be exaggerated, because there is no way to know what pattern of events led to it. The point here, as in the previous analysis, is that these substantial and unaccountable variations indicate the need for a much closer study of the process of education.

In Total Math also there are some puzzling patterns. For example, the gain made by the CON school in Cincinnati was the largest of the four CON school gains, but the gain made by the EXP school in Cincinnati was the smallest of the four EXP school gains. This makes it seem likely that the Math treatment might have been positively effective in all four sites (though perhaps not very large in effect in some of the sites), but for some peculiarity of the situation at Cincinnati. One possible explanation for this anomalous situation in Cincinnati is that the control school regarded the project as a direct challenge. That is, having been told that a new and better idea was to be introduced, they decided to make

(Text resumes on page III-60)

FIGURE III-1

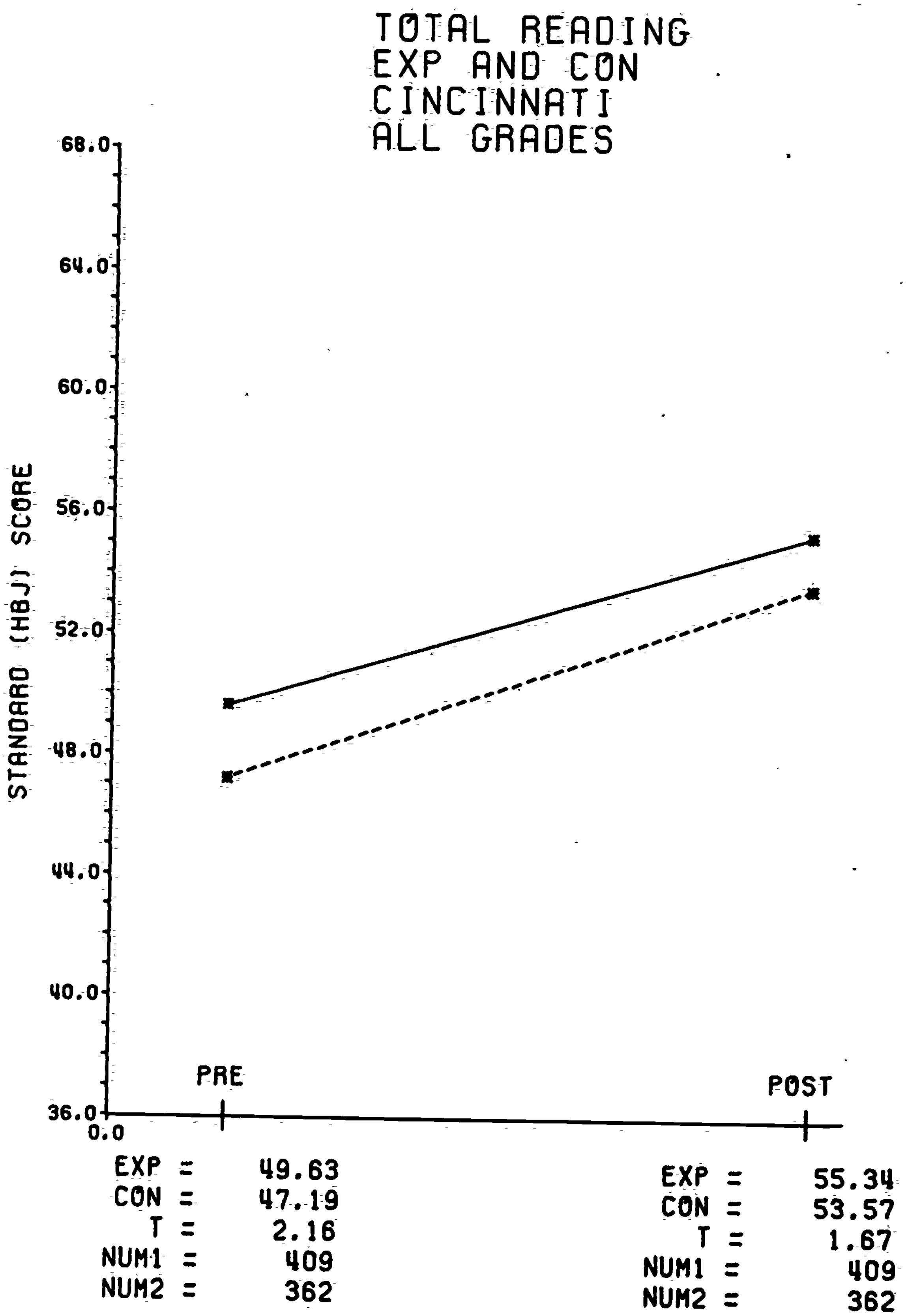
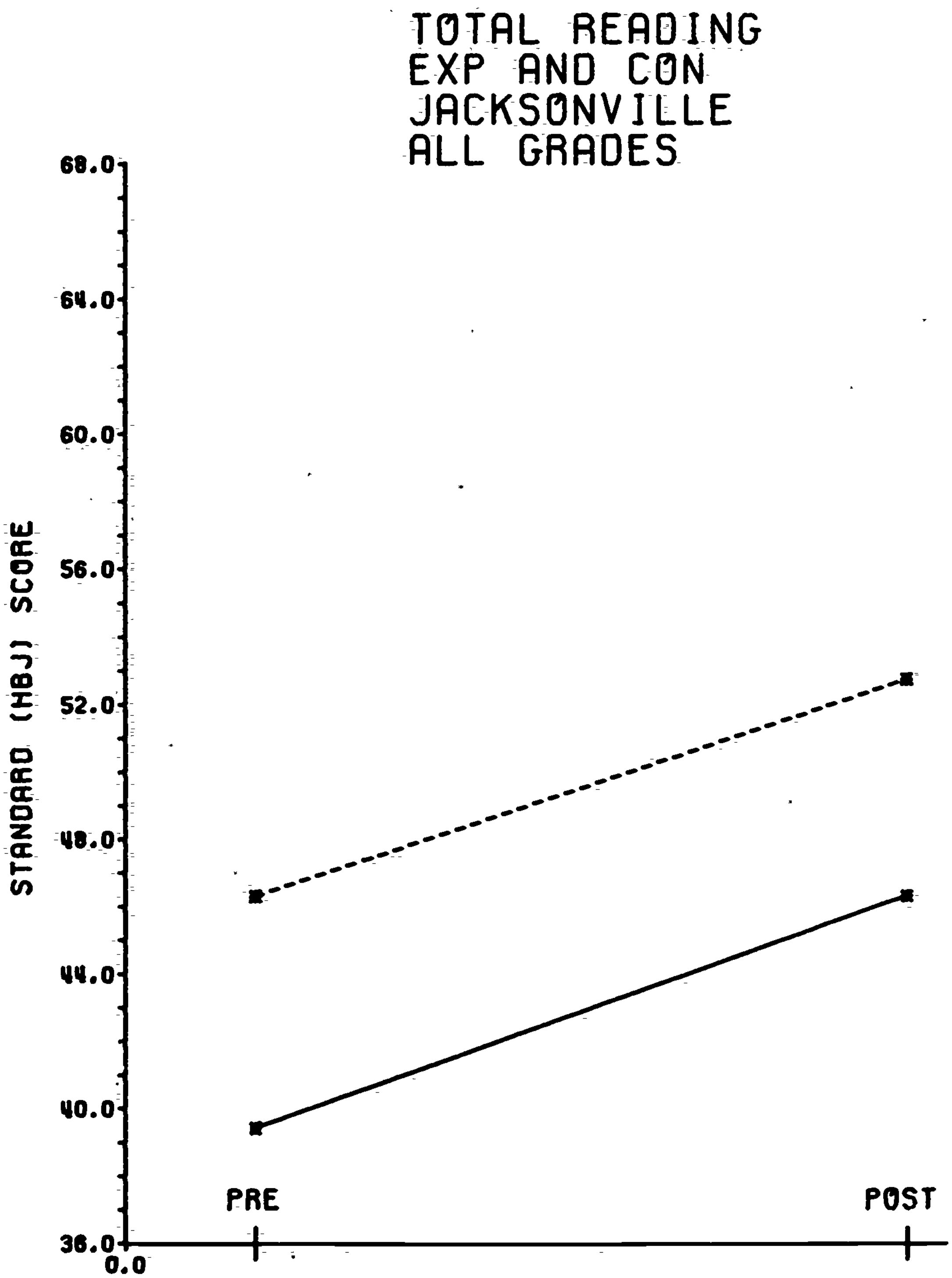
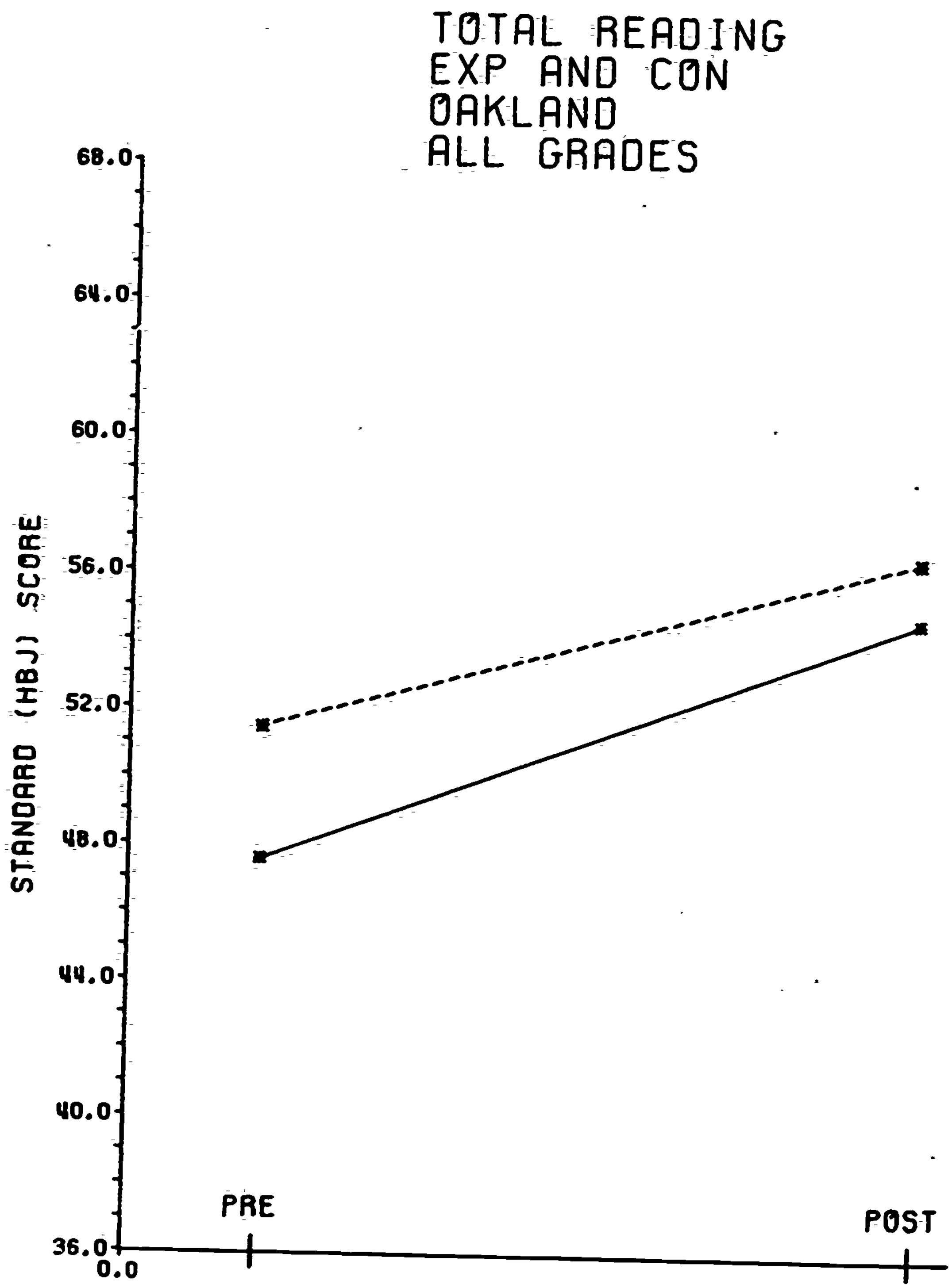

TOTAL READING
EXP AND CON
CINCINNATI
ALL GRADES
68.0
64.0
60.0
56.0
52.0
48.0
44.0
40.0
36.0
0.0
STANDARD (HBJ) SCORE
PRE
POST
EXP = 49.63
CON = 47.19
T = 2.16
NUM1 = 409
NUM2 = 362
EXP = 55.34
CON = 53.57
T = 1.67
NUM1 = 409
NUM2 = 362

FIGURE III-2



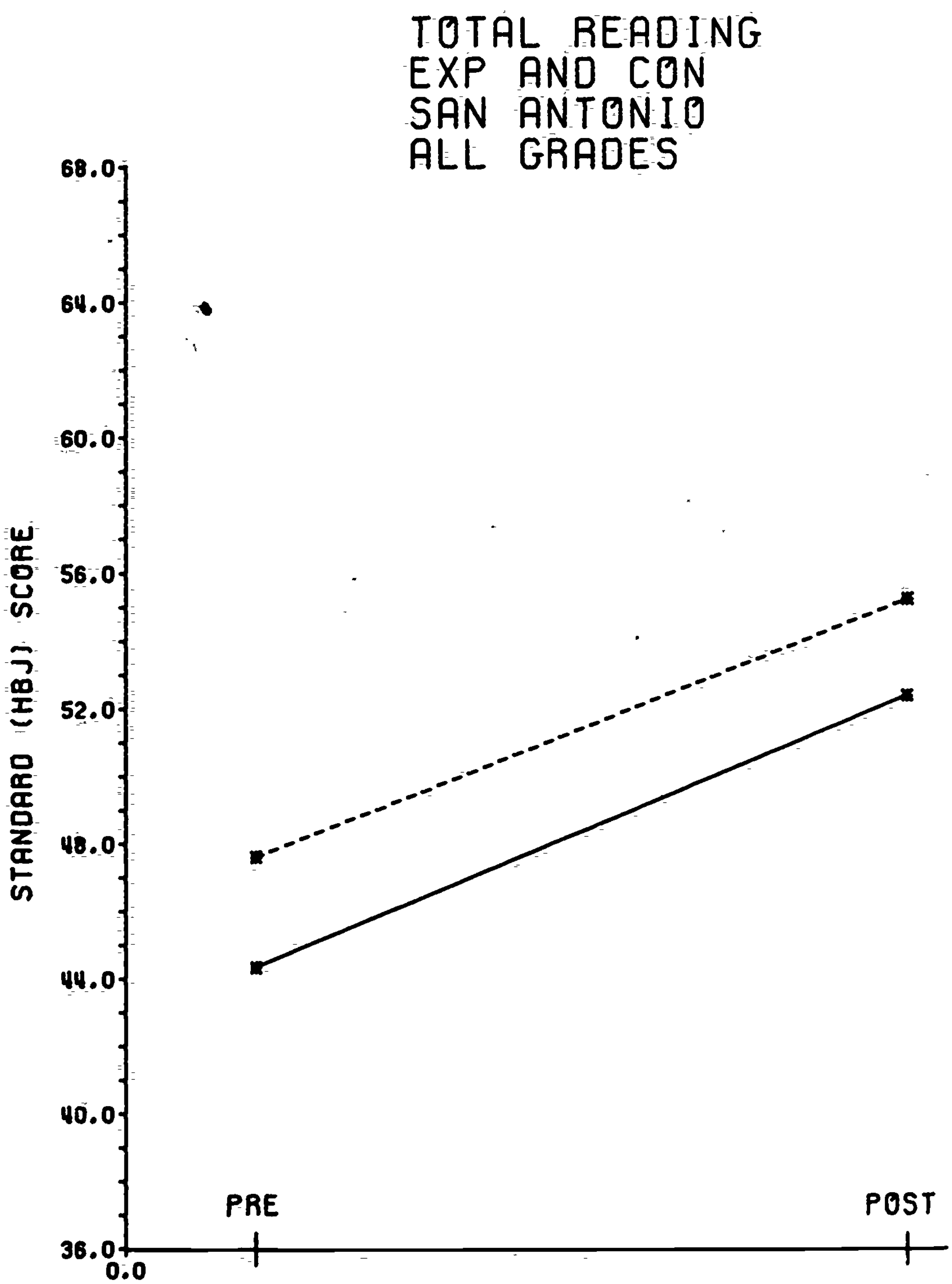
EXP =	39.41	EXP =	46.32
CON =	46.31	CON =	52.75
T =	-8.49	T =	-7.90
NUM1 =	738	NUM1 =	738
NUM2 =	693	NUM2 =	693

FIGURE III-3



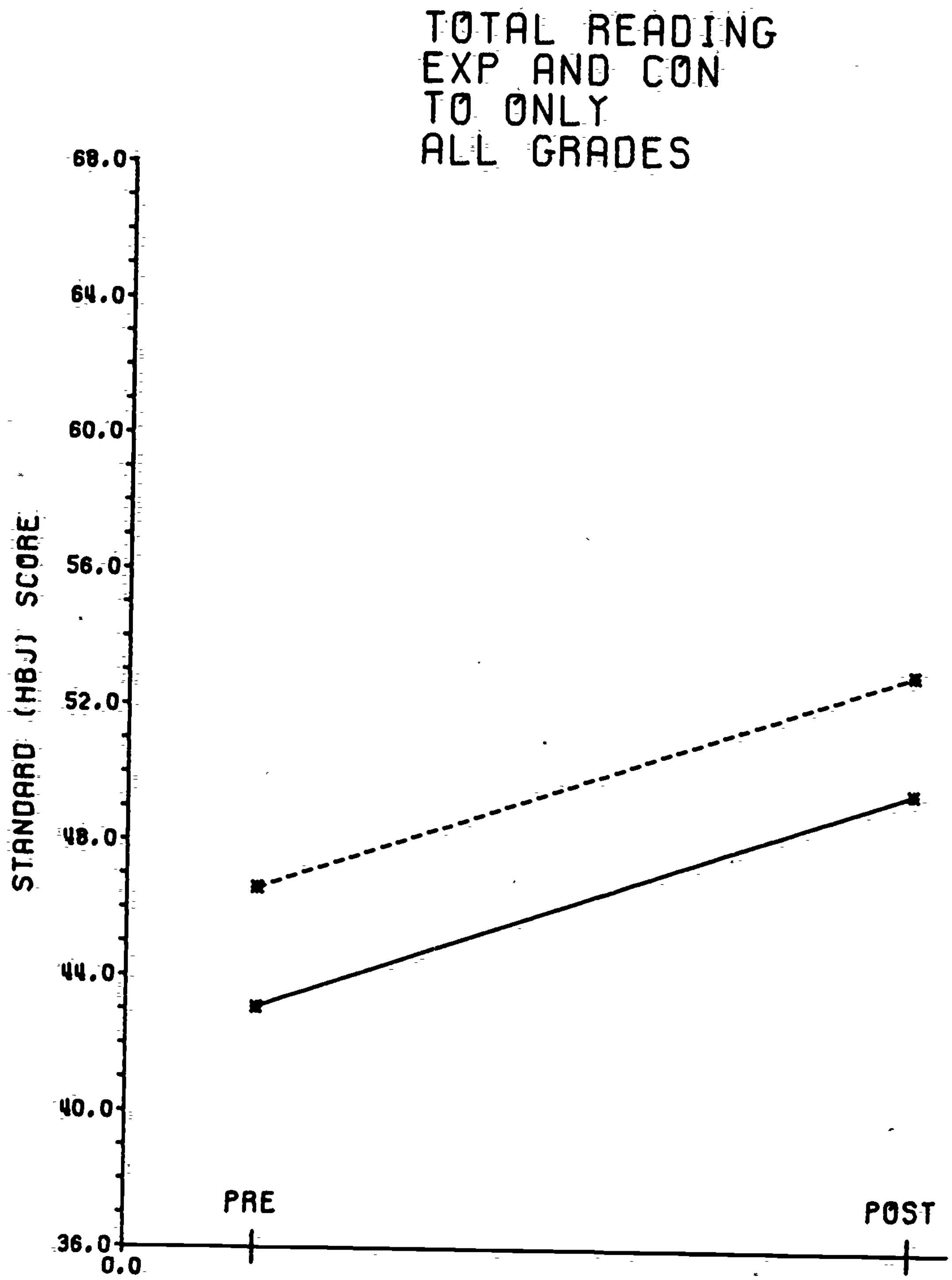
EXP =	47.58	EXP =	54.73
CON =	51.45	CON =	56.49
T =	-3.32	T =	-1.62
NUM1 =	361	NUM1 =	361
NUM2 =	487	NUM2 =	487

FIGURE III-4



EXP =	44.34	EXP =	52.41
CON =	47.62	CON =	55.27
T =	-2.33	T =	-2.16
NUM1 =	333	NUM1 =	333
NUM2 =	288	NUM2 =	288

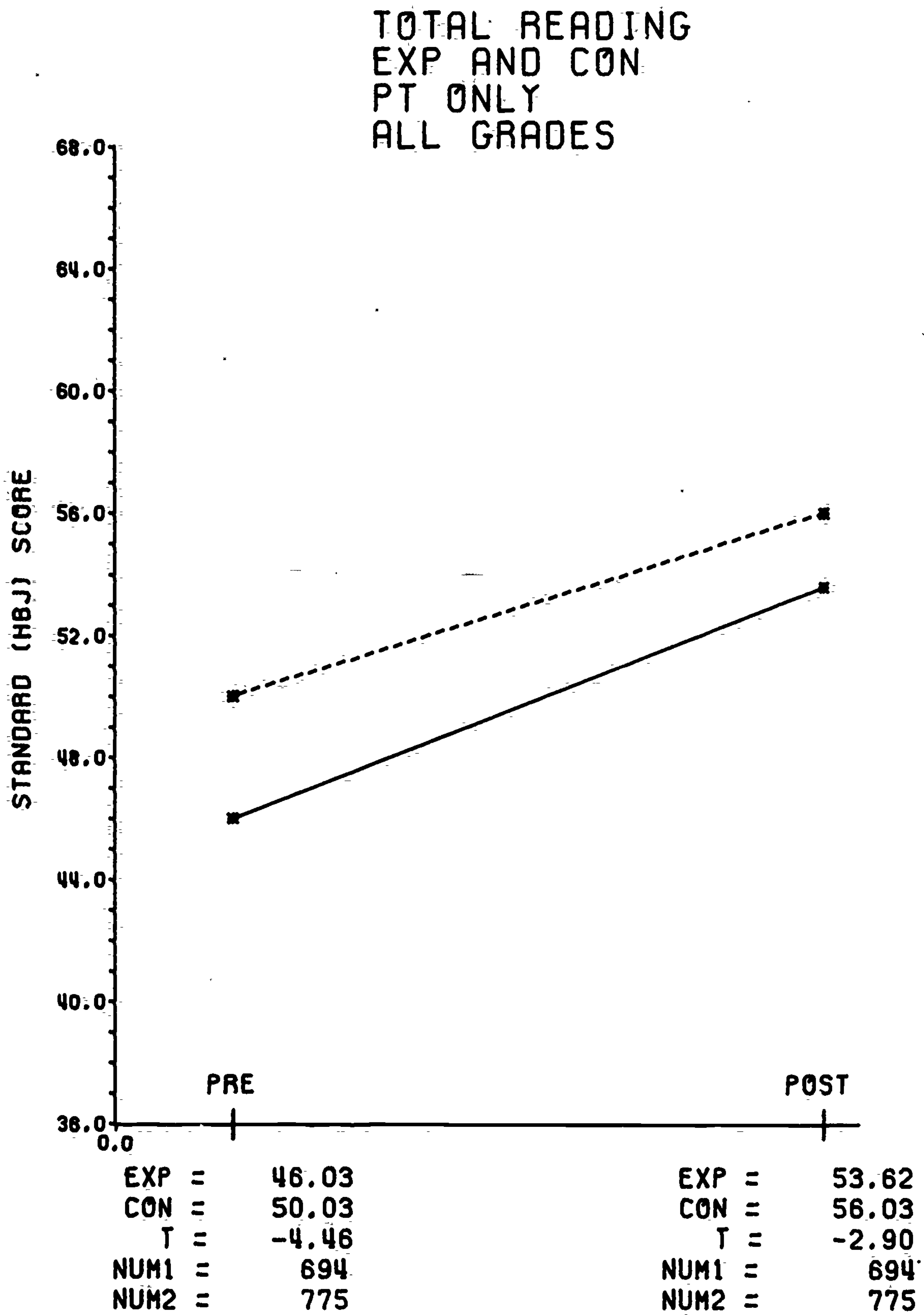
FIGURE III-5



EXP =	43.06
CON =	46.61
T =	-5.26
NUM1 =	1147
NUM2 =	1055

EXP =	49.53
CON =	53.03
T =	-5.31
NUM1 =	1147
NUM2 =	1055

FIGURE III-6



TOTAL READING
EXP AND CON
PT ONLY
ALL GRADES
STANDARD (HBJ) SCORE
68.0
64.0
60.0
56.0
52.0
48.0
44.0
40.0
36.0
0.0
PRE
POST
EXP = 46.03
CON = 50.03
T = -4.46
NUM1 = 694
NUM2 = 775
EXP = 53.62
CON = 56.03
T = -2.90
NUM1 = 694
NUM2 = 775

FIGURE III-7

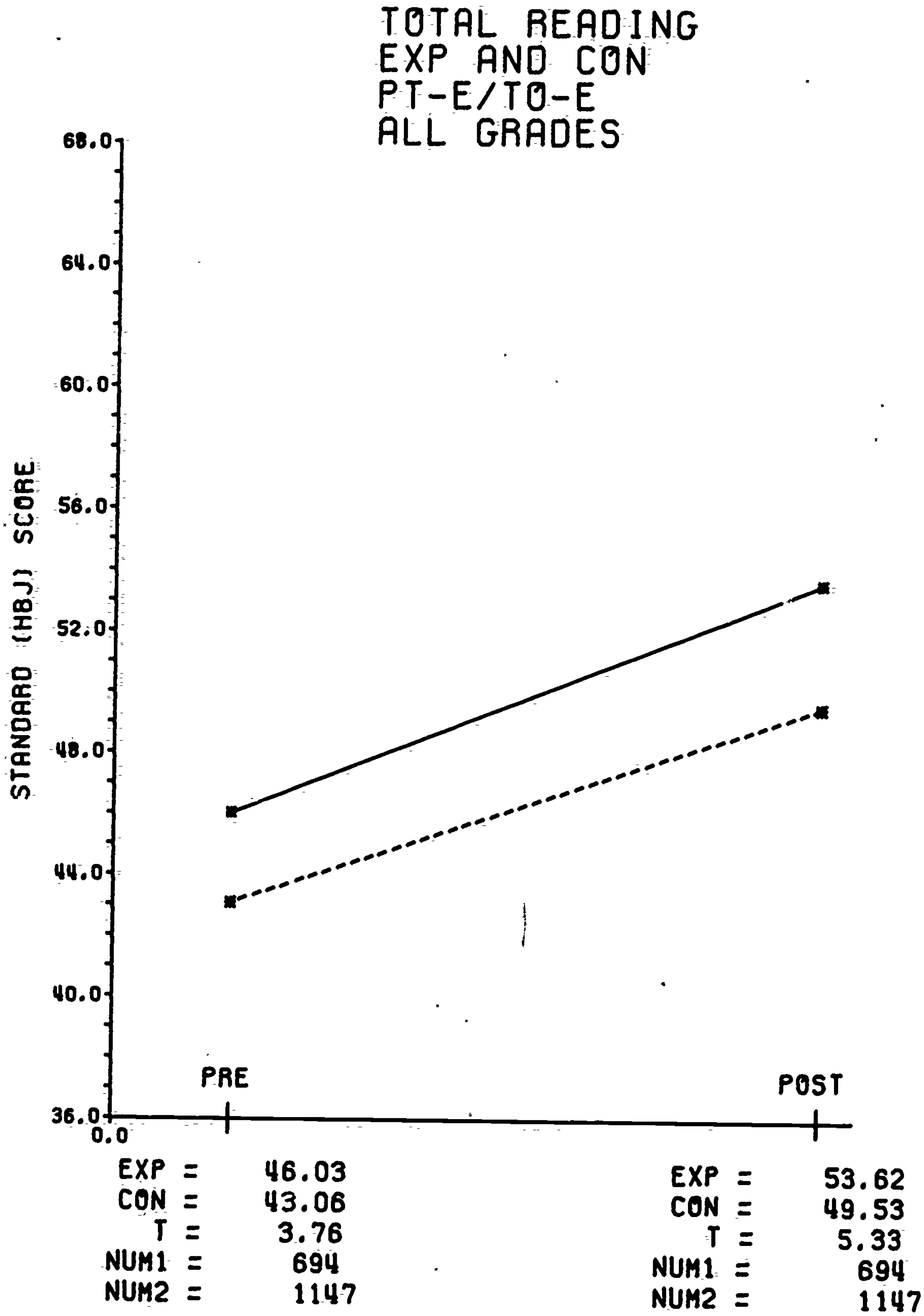

TOTAL READING
EXP AND CON
PT-E/TO-E
ALL GRADES
STANDARD (HBJ) SCORE
68.0
64.0
60.0
56.0
52.0
48.0
44.0
40.0
36.0
0.0
PRE
POST
EXP = 46.03
CON = 43.06
T = 3.76
NUM1 = 694
NUM2 = 1147
EXP = 53.62
CON = 49.53
T = 5.33
NUM1 = 694
NUM2 = 1147

FIGURE III-8

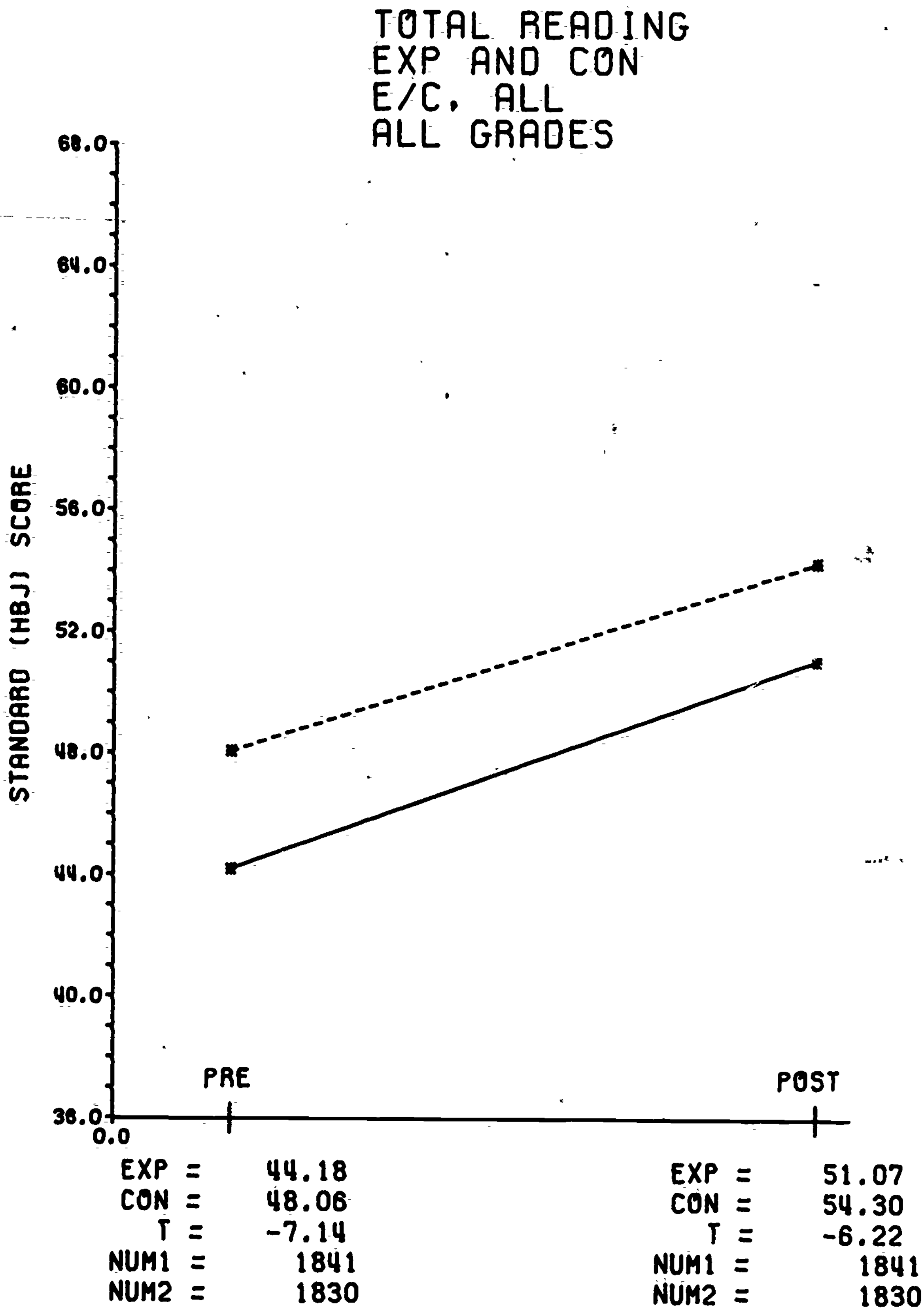

TOTAL READING
EXP AND CON
E/C, ALL
ALL GRADES
68.0
64.0
60.0
56.0
52.0
48.0
44.0
40.0
36.0
0.0
STANDARD (HBJ) SCORE
PRE
POST
EXP = 44.18
CON = 48.06
T = -7.14
NUM1 = 1841
NUM2 = 1830
EXP = 51.07
CON = 54.30
T = -6.22
NUM1 = 1841
NUM2 = 1830

FIGURE III-9

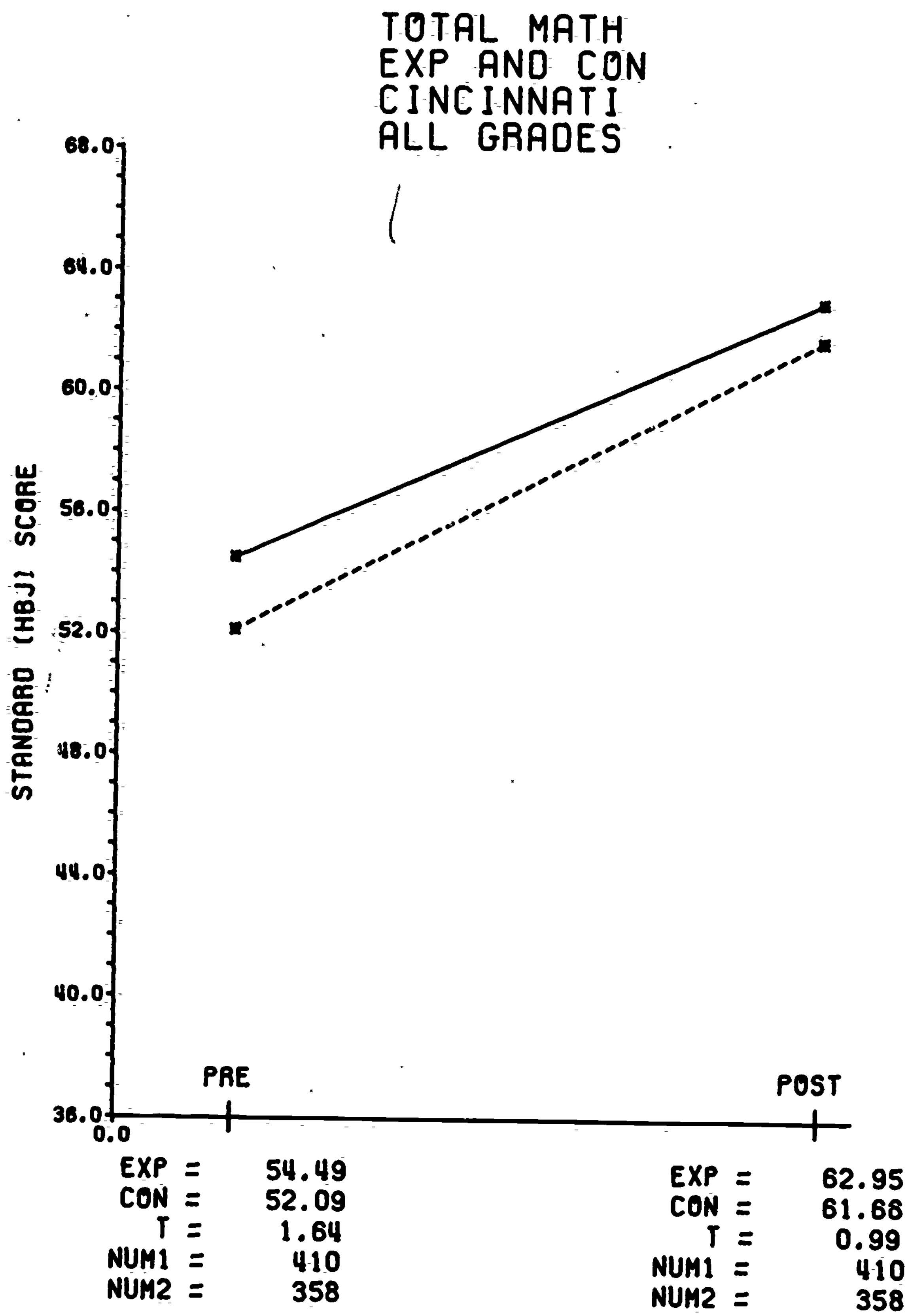
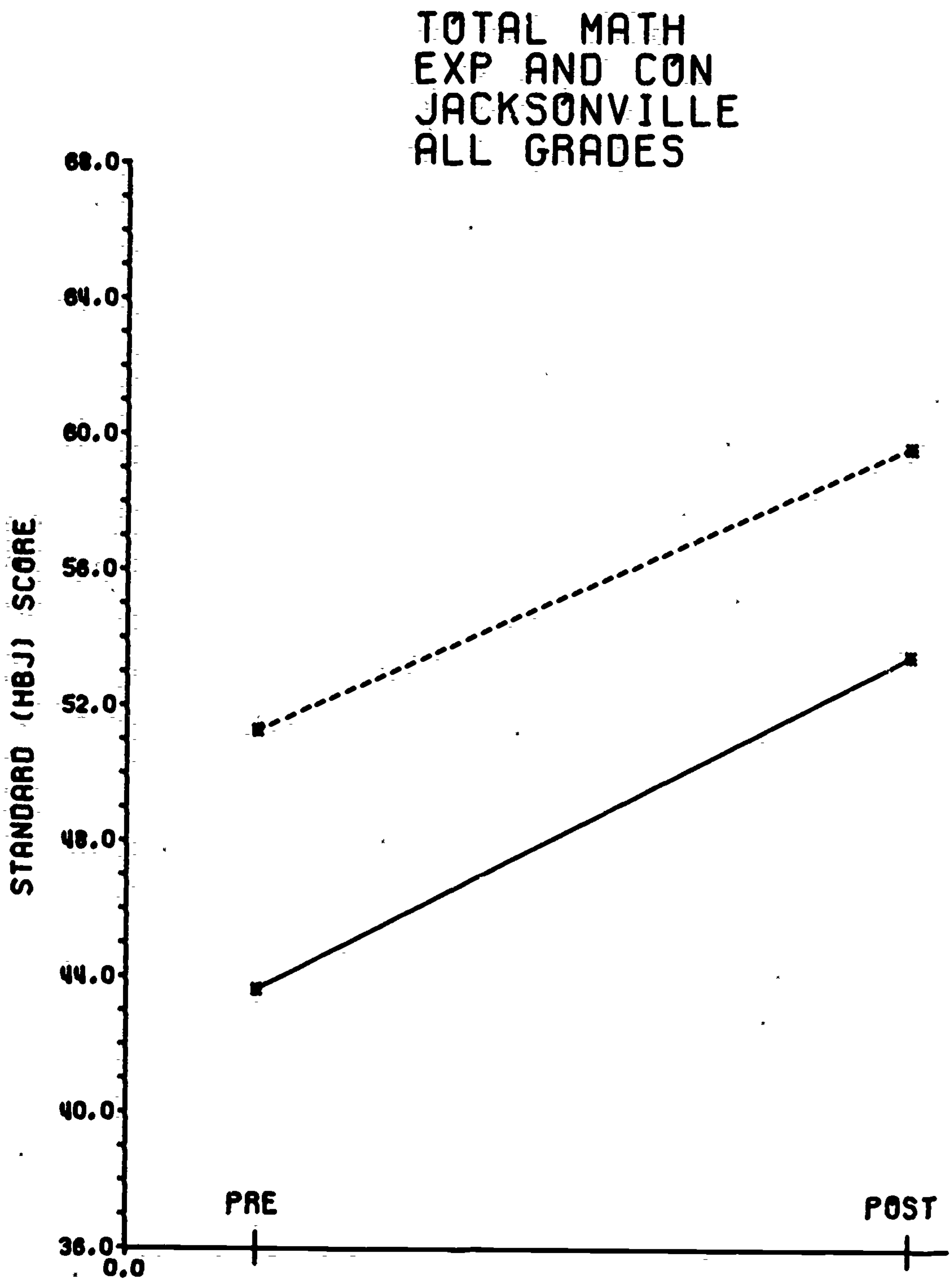
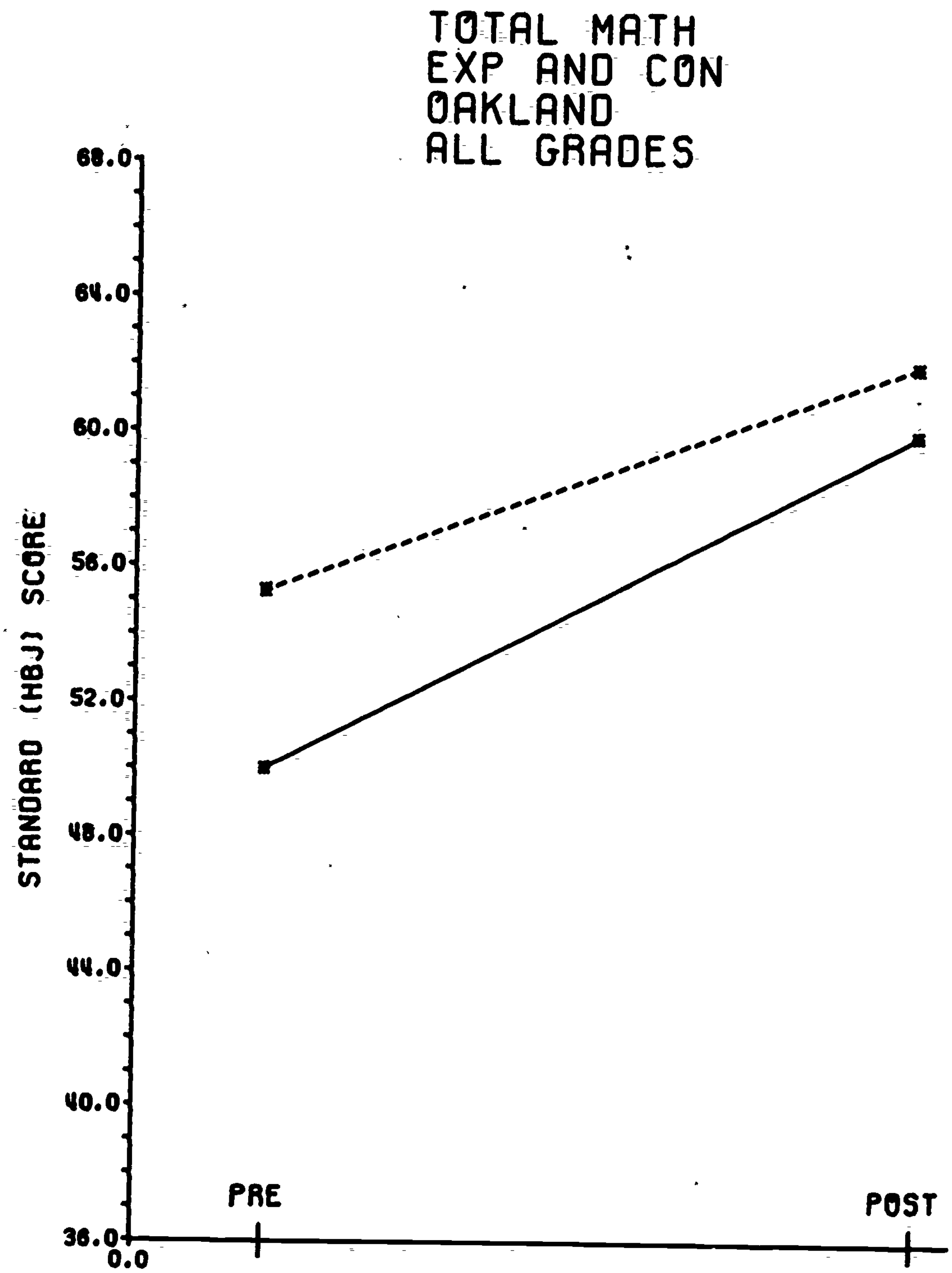

TOTAL MATH
EXP AND CON
CINCINNATI
ALL GRADES
68.0
64.0
60.0
56.0
52.0
48.0
44.0
40.0
36.0
0.0
STANDARD (HBJ) SCORE
PRE
POST
EXP = 54.49
CON = 52.09
T = 1.64
NUM1 = 410
NUM2 = 358
EXP = 62.95
CON = 61.68
T = 0.99
NUM1 = 410
NUM2 = 358

FIGURE III-10



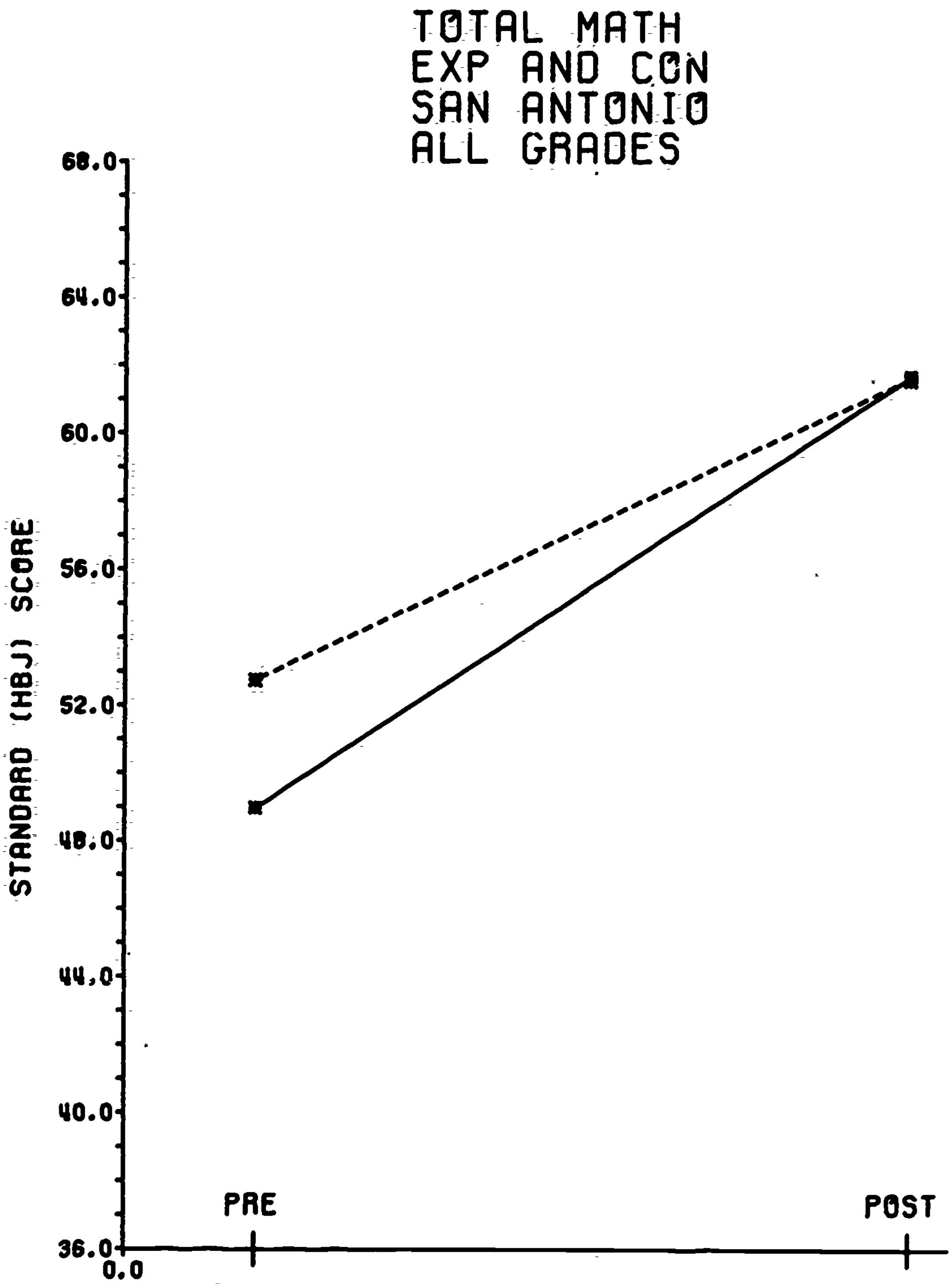
EXP =	43.61	EXP =	53.51
CON =	51.26	CON =	59.63
T =	-6.98	T =	-6.04
NUM1 =	734	NUM1 =	734
NUM2 =	679	NUM2 =	679

FIGURE III-11



EXP =	50.01	EXP =	59.96
CON =	55.27	CON =	61.99
T =	-3.62	T =	-1.60
NUM1 =	357	NUM1 =	357
NUM2 =	488	NUM2 =	488

FIGURE III-12



EXP =	48.98	EXP =	61.70
CON =	52.74	CON =	61.59
T =	-2.21	T =	0.08
NUM1 =	334	NUM1 =	334
NUM2 =	287	NUM2 =	287

FIGURE III-13

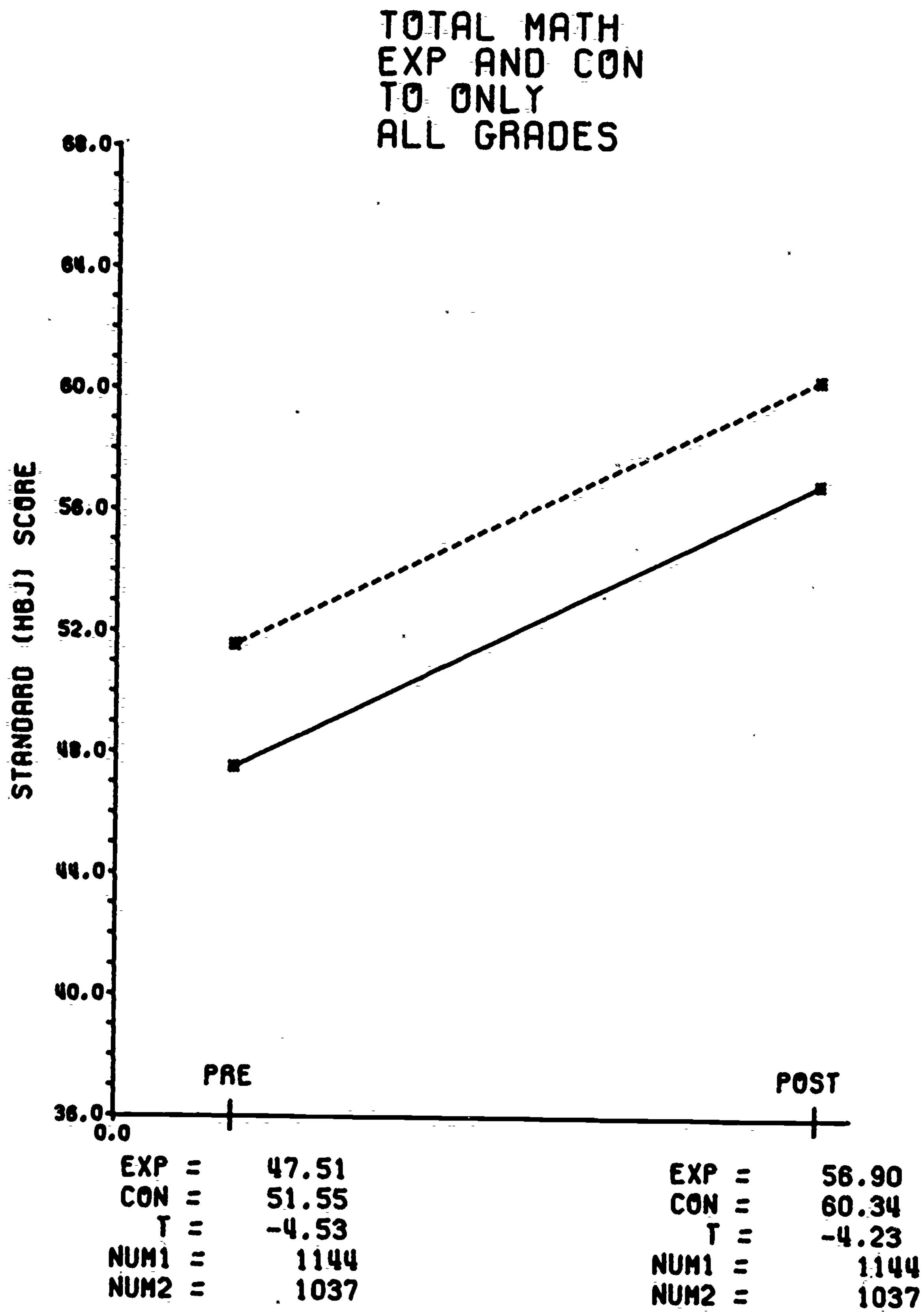
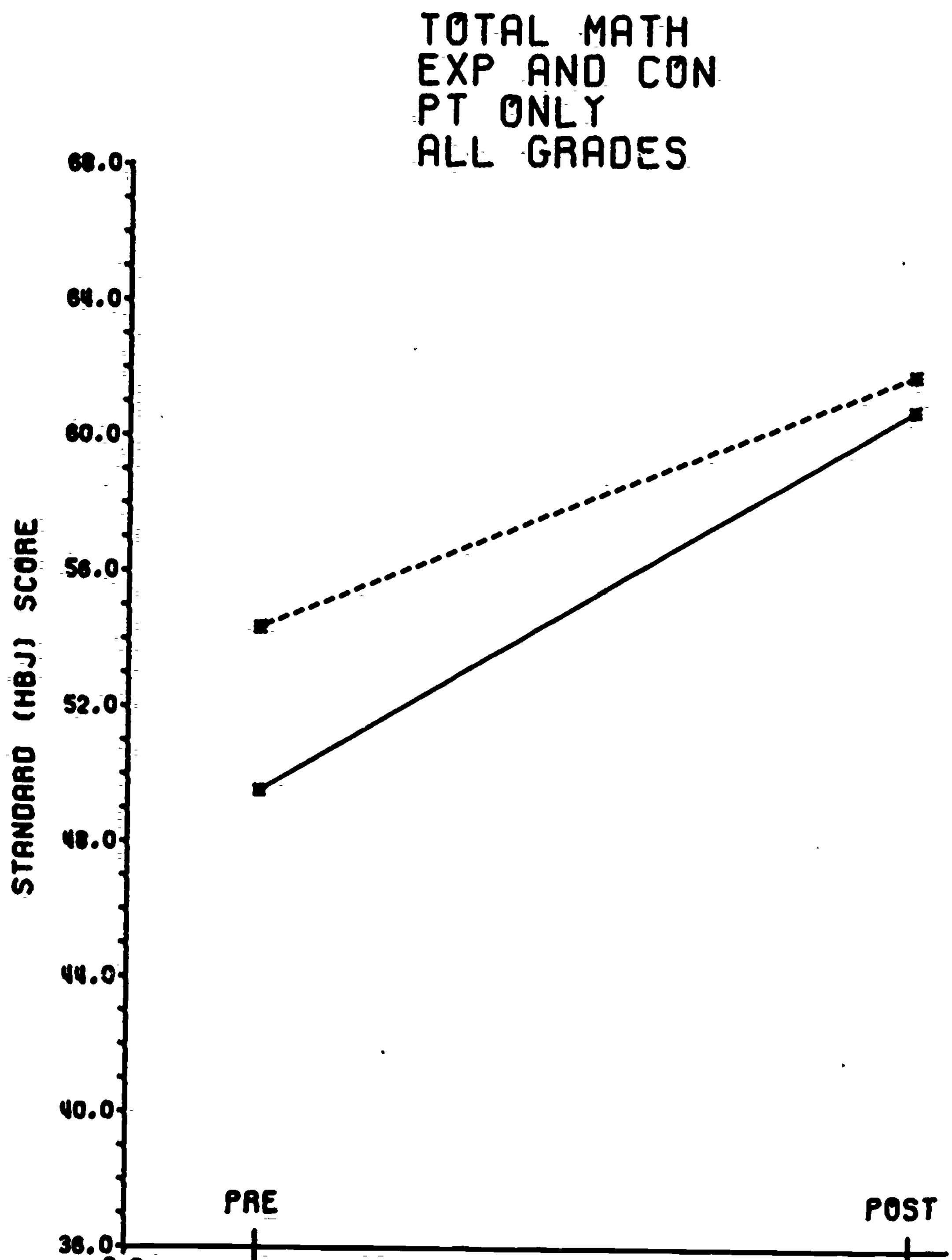
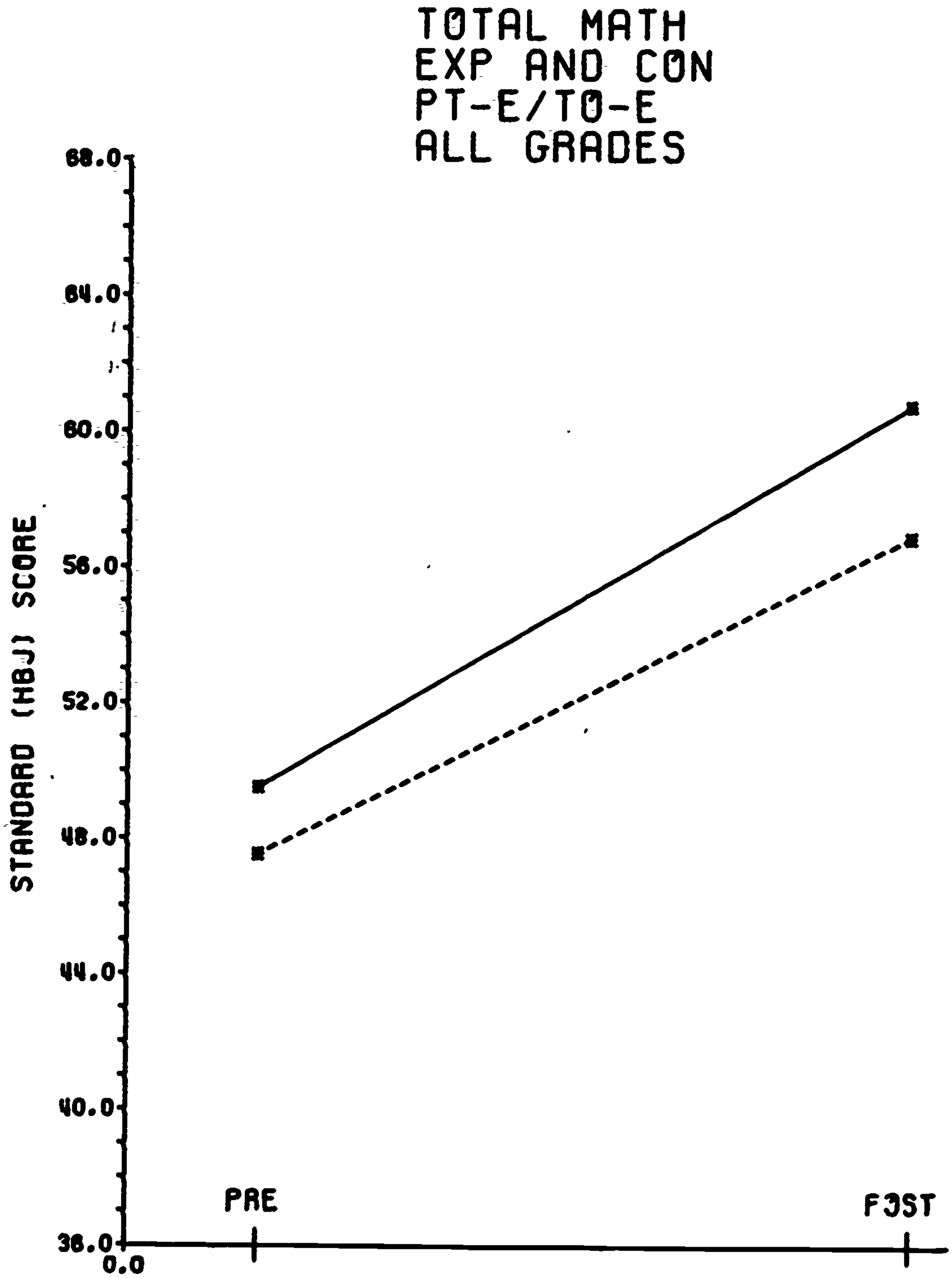


FIGURE III-14



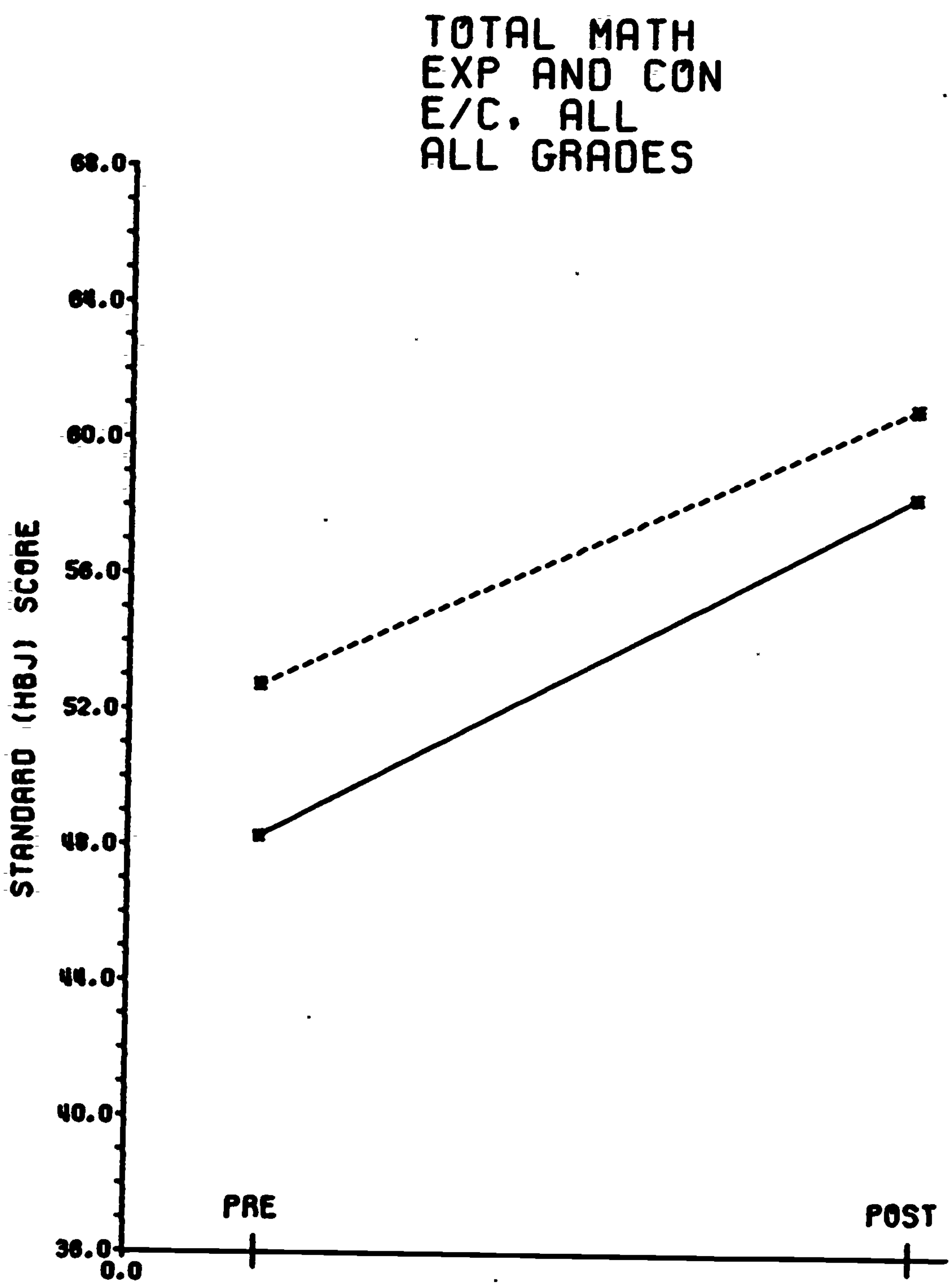
EXP =	49.51	EXP =	60.80
CON =	54.33	CON =	61.84
T =	-4.39	T =	-1.09
NUM1 =	691	NUM1 =	691
NUM2 =	775	NUM2 =	775

FIGURE III-15



	PRE		POST
EXP =	49.51	EXP =	60.80
CON =	47.51	CON =	56.90
T =	2.00	T =	4.33
NUM1 =	691	NUM1 =	691
NUM2 =	1144	NUM2 =	1144

FIGURE III-16



EXP =	48.27	EXP =	58.37
CON =	52.74	CON =	60.98
T =	-6.47	T =	-4.21
NUM1 =	1835	NUM1 =	1835
NUM2 =	1812	NUM2 =	1812

an all-out effort to show that the old way is in fact better than the new one. The results of MAC observations and conversations between TAC staff and staff in the Cincinnati control school indicate that this kind of motivation was operating to some degree there. Thus, there is another complication that must be considered.

Summary

In this section, several sources of possible distortion or ambiguity in the achievement results have been examined. The general conclusions from this examination are that there seem to be no detectable distortions of large magnitude. On the other hand, there are several sources of ambiguity in the achievement results.

Because of this ambiguity, it would be unwarranted to make specific recommendations from this study as to what Federal policy should be with regard to the use of incentives to teachers and parents as a means of improving educational performance. The conclusions which can be drawn from this analysis of the achievement data are of a different nature. The specific conclusions from this further analysis of the achievement data is that although the results must be regarded as tentative, there is a pattern in the achievement data which suggests that incentives may promote increases in learning rate. These increases were most pronounced for the Parent-Teacher model, and for achievement in Math. The pattern does not appear to be due to the methodological artifacts of the analysis made. However, the caution should be made that these results could be due to causes other than the incentives treatment. In fact, the analysis of the attitudinal and behavioral data from students, teachers, and parents, which is presented in the subsequent sections, will tend to weaken the argument that the observed achievement patterns are simply the results of the incentives treatment, because there will be very little consistency between the achievement results and the results of these other analyses.

A second, more general, methodological conclusion emerged from this further analysis of the achievement data. There are several analytical issues involved in the use of achievement test data for this kind of

purpose. The possibility of floor and ceiling effects on test scores was recognized by the project designers, and the use of the multilevel test administration system was an attempt to deal with it. Similarly, the use of carefully trained and supervised test administrators and explicit procedural instructions was an attempt to minimize the potential lack of reliability and differential reliabilities. However, these testing procedures have not been empirically validated with these data; therefore, at this point these are no more than assumptions about how tests ought to be assigned and administered. Furthermore, additional technical questions remain: the choice of the most accurate score format in multilevel testing; the more general problem of regression effects of school means; and the fan-spread phenomenon. The latter two issues have only recently been recognized, and far more research into them must be carried out if educational research is to inform policy makers wisely.

Analysis of Student Attitudes and Behavior

Introduction

The analysis of the questionnaire, interview, and observation data presents somewhat different challenges than those posed by the achievement data. For example, with the achievement data, it was necessary to make some decision as to score format, but the content of the scale was established by the publisher. For the questionnaire, interview, attendance, and observation data, however, the decisions as to index construction had to be made during the analysis. Circumstances of the project, particularly the very tight time schedule and the problems of obtaining clearances for the instruments, coupled with the fact that the unusual nature of the project made it possible to adopt existing instruments, precluded any full scale program for the validation of the survey instruments. Instead, special *ad hoc* instruments were developed which asked the relevant questions in a direct and simple way.

The content of these instruments was determined by the hypotheses that had been formulated concerning the potential impacts of the incentives models. For example, one group of hypotheses asserted that students in the EXP schools would become more favorable in attitude toward several aspects of the school experience. In particular, it was hypothesized that students in the EXP schools would become more favorable toward (1) reading school work, (2) arithmetic school work, (3) teachers, and (4) school in general. An additional hypothesis was that students in the EXP schools would become more optimistic and confident about further school experiences, and would develop stronger aspirations for high educational attainment. Also, it was hypothesized that students in the EXP schools would become more concerned about the performance of their fellow students, and more favorably disposed to assist them. Finally, it was hypothesized that students in the EXP schools would develop more favorable self-images and perceptions of their achievements and opportunities. This last is a deliberately broad concept, embracing some elements in the concept of orientation toward future schooling as well as

some direct self-image elements. These hypotheses led to the development of items for the questionnaires and interviews which would tap each variable described above. Such items were included both in the Student Questionnaire and Student Interview Schedule. A similar procedure was followed in developing the data collection instruments for use with teachers and parents.

The instruments were designed to be easy to administer. As a result of these efforts, and the careful planning for the administration, the completion rate for items within questionnaires and interviews for this project is much higher than the usual level encountered in studies dealing with disadvantaged elementary students. However, another consequence of the necessary compromises in the development procedure is that the instruments have only face validity (which is quite high, given the decision to make the questions simple and direct), and have unknown reliability.

A second issue with respect to the attitude and behavior data, and one which applied to parents and teachers as well as to students, concerns the timing of the first wave of data collections. For the student interviews and the classroom observations, the first wave of data was collected from the middle of November to the end of December. Although this was substantially after the beginning of the school year, it can still be regarded with some assurance as a "pre" measure, since the data were collected before the incentives models had exerted any major impacts on the schools. In fact, the local districts did not actually sign grants to participate in the study until January, and the parents and teachers did not have contracts to sign until February and March. All of the data from the interviews (student, teacher, and parent) had been collected prior to that time.

However, for the student (and also the parent and teacher) questionnaires, the first wave of data was not collected until late March, which is quite late in the year. The problem that this creates for analysis is that it thus becomes quite problematic whether this administration should be regarded as being <u>before</u> the treatment, or <u>during</u> it. In other words,

although there are two waves of questionnaire data, it is not at all clear that they can be analyzed as a before-after comparison. For example, if the data were to be considered as reflecting the before and after conditions, and analyzed that way, and if no differences (impacts of the treatment) were found, what conclusion could be drawn? One analyst might conclude that the treatment did not have any impact, and another might decide that perhaps after all the impact had occurred before the first (March) measurement.

Obviously, in this situation, the interpretation of each result must be made somewhat subjectively, and on the basis of the relevant data for each scale and each comparison. Consequently, the results are presented in tables that show, for every index, the average of the EXP and CON at the first administration, and the average of the EXP and CON at the second administration.

Even under these difficult circumstances, however, there are some possible patterns of outcome which seem unambiguous. For example, if there is a difference favoring the experimental group at the second administration, but a considerably smaller difference at the first administration, then (except for possible regression effects) one can rather safely conclude the incentives treatment had an impact which occurred during the interval between the first and second administration of the instrument. If, on the other hand, there is a difference on the second administration favoring the EXP group, but that same difference exists at the first administration, the conclusion cannot be definite. The analyst (and the reader) must then decide how likely it is that the difference found at the first administration represents a pre-existing difference, and how likely it is that it represents instead a quick impact of the incentives treatment. Since the schools within a site were not assigned at random to an EXP or CON treatment, there is no way of arguing that any pre-existing differences will be dampened out when sites are combined. The only guideline available in this situation is that one must be consistent in going from item to item or index to index. If the conclusions drawn from the results of one index are inconsistent with the

conclusions drawn from the results of another index, then the only prudent reaction is to suspend judgment.

If there are no differences at the second administration favoring the EXP group, or perhaps even a difference favoring the CON group, then again the situation may be indeterminate. The results of all this analysis, in short, are dependent upon the judgment of the reader. This is an unavoidable consequence of the design and execution of the study. These results should thus be regarded as providing not firm knowledge about the impacts of the incentives models, but rather as indicating promising areas for future investigation.

Measuring the practical magnitude of an impact under these conditions is of course doubly difficult, since even the numerical value of the impact is not certain. However, in instances where there is some reason for confidence in the real existence of an impact, the size of that impact in substantive (educational) terms will be estimated as closely as possible.

Regression of scores toward a common mean is more likely with these survey items than it was with the achievement test results, since these items are used individually or in indexes based on only a few items, rather than being scores based on the sum of a large number of items. In other words, unreliability is more of a problem with these survey items than it was with the achievement test scores.

As with the achievement results, there are always Hawthorne effects, John Henry effects (Saretsky, 1972) and possibly the opposite of a John Henry effect, what might be called the Rosenthal effect (Rosenthal,1966). There is, however, no particular reason to expect a fan-spread phenomenon to be occurring on these survey items.

Index Construction

Data on student attitudes and behavior were collected from four sources, at several points in time. The sources included:

1. A student questionnaire which all 4600 students were asked to complete

2. A student interview which was completed by a random sample of approximately 200 students

3. Attendance records for each classroom each month

4. Classroom observations on student behavior, collected within each of the approximately 160 classrooms of the project.

Because all of this data provides information about the impacts on students as a result of the incentives project, all of it will be discussed in this section. Subsequent sections discuss the parent and teacher results.

For the Student Questionnaire, there were 43 items. These were presented in such a way that reading skill was not essential to completing the task. The first 13 items were coded as 1 or 2 or 3, with the higher number always associated with the response that appeared most desirable. The remaining items were coded 1 or 2, with the higher number again always associated with the response that appeared to be most desirable. Thus, the indexes created by addition of responses always were oriented so that a higher score meant a more favorable outcome. From these 43 items, seven indexes were constructed, based on the research hypotheses and the item content. These indexes were created by adding together the responses to each of the component items. If a response to any item was missing for a particular student, then the value of the index also was considered to be missing for that student. Since the completion rate for items on the Student Questionnaire and interview was very high (averaging over 95%), it was possible to follow this simple and desirable strategy without losing a substantial amount of data. The names of the indexes constructed from the Student Questionnaire items, and the items which constitute each, are shown in Table III-12.

For the Student Interview, there were 15 items. From these 15 items six indexes were constructed. The indexes parallel those developed from the Student Questionnaire; however, there was no index of "Orientation

TABLE III-12

INDEXES CONSTRUCTED FROM THE STUDENT QUESTIONNAIRE

1. Student Attitude Toward Reading

Question No.	
1	When I think about learning to read
3	When I think about reading
5	When I think of my reading teacher
15	Reads well vs does not read well
18	Likes to read at school vs does not like to read at school
20	Reads alot at home for fun vs watches TV or plays
21	Goes to a library to read for fun vs does not go to a library

2. Student Attitude Toward Arithmetic

Question No.	
2	When I think about learning to do arithmetic problems
4	When I think about doing arithmetic problems
6	When I think of my arithmetic teacher
14	Good at doing arithmetic problems vs not good at
17	Likes to work at school with arithmetic problems vs does not
22	Does arithmetic problems at home for fun vs watches TV or plays
28	Wants to be good at doing arithmetic problems vs does not care

3. Student Attitude Toward Teachers

Question No.	
5	When I think of my reading teacher
6	When I think of my arithmetic teacher
8	When I talk to my teacher in school, the teacher looks
10	When one of the kids asks a question, the teacher looks
12	When we try to have fun in class, our teacher looks

4. Student Attitude Toward School

Question No.	
13	Goes to a good school vs goes to a bad school
16	Likes to go to school vs does not like

5. Student Orientation Toward Future Schooling

Question No.	
27	Wants to do well in school vs does not care how well he does
33	Expects to go to high school vs does not expect
38	Wants to go to high school vs does not want to go
39	Wants to get training for a good job vs does not care
41	Expects to be trained for a good job vs does not expect to be

- Continued

TABLE III-12 (Cont'd.)

6. Student Relations with Other Students

Question No.	
9	When other kids ask me for help in their school work, I feel
11	When another kid in my class makes a mistake, I feel
24	Help other kids outside of school with their schoolwork vs does not
25	Helps other kids at school with their school work vs does not
26	Wants all the kids to do well in school vs does not care

7. Student Orientation Toward Self

Question No.	
14	Good at doing arithmetic problems vs not good at
15	Reads well vs does not read well
27	Wants to do well in school vs does not care how well he does
28	Wants to be good at doing arithmetic problems vs does not care
29	Wants to be a good reader vs does not care
30	Most people think child is good vs most people think child is bad
31	Expects to be a good reader vs does not expect to be
32	Easy to learn things vs hard to learn things
33	Expects to go to high school vs does not expect to go to high school
34	Can do the things he wants vs cannot do the things he wants
35	Expects to be good in arithmetic vs does not expect to be good
36	Finishes his work vs gives up easily on his work
37	Likes himself most of the time vs does not like himself
40	Finishes his work vs will get what he wants because of luck
42	Will get the job he wants vs will not get the job he wants

toward Self" developed from the Student Interview data. The same procedures were followed in constructing the Student Interview indexes as had been used in constructing the indexes for the Student Questionnaire data. It should be noted, however, that the number of cases for the questionnaire responses is approximately 3800, and the number of cases for the interview responses is approximately 200. The names of the indexes for the Student Interview items, and the items from which each was constructed are shown in Table III-13.

Because the content of these two instruments is similar, and because the indexes refer to attitudinal rather than behavioral dimensions, it seems best to discuss together the results from the Student Questionnaire and Student Interview data. After that discussion is completed, a discussion of the attendance data and the classroom observation data on student behavior will follow this presentation.

Single-Site Comparisons--Student Questionnaire and Interview Results

The first comparison examines the results of the Student Questionnaire and Student Interview data at Cincinnati, and compares the EXP and CON schools there. Cincinnati is one of the sites at which the Teacher-Only (TO) model was used. The results for the Time 1 and Time 2 administrations are presented in Tables III-14 and III-15, respectively. The general picture in the Cincinnati data on student attitudes is fairly clear. For Cincinnati, there are few changes in t-value between the first and second administration of the instrument that are worthy of note. Interestingly, the few changes (*e.g.*, in "Attitude toward Arithmetic" and "Relations with Other Students" from the Student Questionnaire, and in "Orientation toward Future Schooling" from the Student Interview) are in the direction that indicates change toward <u>less</u> favorable attitudes among students in the EXP schools. However, two of these changes (those from the Student Questionnaire) might be due to regression effects, since students in the EXP school were initially very high on these indexes. The t-value for the Student Interview index of "Orientation toward Future Schooling" at

TABLE III-13

INDEXES CONSTRUCTED FROM THE STUDENT INTERVIEW

1. Student Attitude Toward Reading

Question No.	
4	Do you enjoy work in reading, or would you rather not do it
5	How many times do you visit the library on your own, out of class
9	Would you like to be a good reader
13	Do you expect to be a good reader when you leave school

2. Student Attitude Toward Arithmetic

Question No.	
3	How do you feel about doing work in arithmetic, do you enjoy it or would you rather not do it
8	Would you like to be good at doing arithmetic or don't you care
14	Do you expect to be good at arithmetic when you leave school

3. Student Attitude Toward Teachers

Question No.	
11	Does your teacher give you a hard time if your work is not good
15	How do you get along with your teacher, well, or not so well

4. Student Attitude Toward School

Question No.	
1	Do you enjoy school. Tell me how you feel in the morning before coming to school--glad or not so glad
2	When you get to school, do you like the work you do in class

5. Student Orientation Toward Future Schooling

Question No.	
10	When you finish school, do you expect to get a good job

6. Student Relations with Other Students

Question No.	
6	When you do your school work, do you get any help from your friends
7	Do you help your friends with their school work

TABLE III-14

STUDENT QUESTIONNAIRE AND INTERVIEW RESULTS
CINCINNATI, TIME 1

INDEX	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\bar{X}$	S.D.	N	$\bar{X}$	S.D.	N	t-val.	r_{pb}
SQ READ	14.61	1.82	474	14.83	1.64	400	-1.83	-.062
SQ ARITH	15.08	1.85	479	14.42	2.36	401	+4.70	+.113
SQ TEACHER	12.15	2.12	478	12.41	2.28	405	-1.78	-.060
SQ SCHOOL	3.61	0.64	478	3.63	0.59	404	-0.38	-.014
SQ FUTURE SCHOOL	9.67	0.68	474	9.61	0.61	401	+1.13	+.037
SQ REL. W/OTHERS	10.15	1.47	478	9.72	1.61	400	+4.14	+.139
SQ SELF	29.17	2.05	472	29.15	2.23	398	+0.15	+.000
SI READ	15.71	1.57	24	16.33	1.83	21	-1.23	-.185
SI ARITH	13.71	2.12	24	14.29	1.45	21	-1.05	-.158
SI TEACHER	7.92	1.86	24	8.14	1.56	21	-0.44	-.066
SI SCHOOL	9.00	1.38	24	9.14	1.59	21	-0.32	-.049
SI FUTURE SCHOOL	4.96	0.20	24	4.95	0.22	21	+0.09	+.014
SI REL. W/OTHERS	4.13	2.13	24	4.29	2.37	21	-0.24	-.036

TABLE III-15

STUDENT QUESTIONNAIRES AND INTERVIEW RESULTS
CINCINNATI, TIME 2

INDEX	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\bar{X}$	S.D.	N	$\bar{X}$	S.D.	N	t-val.	r_{pb}
SQ READ	14.40	1.90	451	14.58	1.99	391	-1.35	-.047
SQ ARITH	14.84	2.05	451	14.60	2.29	387	+1.56	+.054
SQ TEACHER	11.54	2.31	458	11.93	2.45	393	-2.37	-.081
SQ SCHOOL	3.39	0.73	458	3.42	0.73	394	-0.61	-.020
SQ FUTURE SCHOOL	9.65	0.69	456	9.55	0.92	387	+1.82	+.062
SQ REL. W/OTHERS	9.73	1.66	453	9.51	1.61	390	+1.95	+.067
SQ SELF	28.81	2.48	433	28.75	2.39	371	+0.34	+.010
SI READ	15.71	2.05	17	16.67	1.50	18	-1.59	-.266
SI ARITH	13.76	1.60	17	14.50	1.26	16	-1.46	-.253
SI TEACHER	7.94	1.95	18	8.56	1.42	18	-1.07	-.181
SI SCHOOL	8.44	1.62	18	9.06	1.39	18	-1.22	-.204
SI FUTURE SCHOOL	4.72	0.96	18	5.00	0.00	18	UNDEF	UNDEF
SI REL. W/OTHERS	4.67	2.11	18	5.06	2.44	18	-0.51	-.087

Time 2 is not calculable, since there was no within-group variance on this index for the CON school. However, the difference between the EXP mean and the CON mean at Time 1 is +.0059. The difference between these two means at Time 2 is -.2778. Thus, the Time 2 difference is in the direction favoring the CON group, and is not negligible in size. For Cincinnati generally, as far as students self-expressed attitudes and orientations are concerned, there is only one t-value above 2.00 in magnitude. This final t-value greater than 2.00 occurs on the Student Questionnair index "Attitude toward Teachers" presented in Table III-15, Time 2. It indicates a <u>less</u> favorable attitude toward teachers exists in the EXP school than in the CON school at the second time. The same direction was also observed when the first wave questionnaire was administered. This pattern is opposite in direction to any possible regression effect.

Turning next to Jacksonville, one finds a different and somewhat puzzling pattern. Here, four of the Questionnaire-based indexes show changes that seem to indicate favorable impact of the treatment. The treatment used at Jacksonville was the Teacher-Only model. The results for Jacksonville are presented in Tables III-16 and III-17. The four indexes from the Student Questionnaire which show changes in the predicted direction are (1) "Attitude toward Arithmetic", (2) "Attitude toward Teacher", (3) "Attitude toward School", and (4) "Relations with Other Students". Another index, "Attitude toward Reading", tends in the same direction, although less strongly than the first four mentioned. However, on two other Questionnaire-based indexes, namely "Orientation toward Future Schooling" and "Orientation toward Self", there is a clear change favoring the students in the CON schools. In the cases of "Orientation toward Future Schooling", there is no difference at the first administration, but a difference favoring the CON schools appears at Time 2. In the case of the "Orientation toward Self", there is a difference favoring the EXP schools at Time 1, but that difference has disappeared at Time 2. This latter pattern could be due to regression toward the mean. It should be noted also that the EXP school in Jacksonville has, judging from the Student Questionnaire data, a more favorable general climate among its

TABLE III-16

STUDENT QUESTIONNAIRE AND INTERVIEW RESULTS
JACKSONVILLE, TIME 1

INDEX	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
SQ READ	15.19	1.59	759	14.77	1.63	586	+4.76	+.129
SQ ARITH	15.37	1.71	762	14.78	2.01	587	+5.82	+.157
SQ TEACHER	12.77	2.06	773	11.95	2.14	591	+7.21	+.192
SQ SCHOOL	3.86	0.41	783	3.76	0.52	593	+3.91	+.105
SQ FUTURE SCHOOL	9.62	0.76	755	9.60	0.75	586	+0.47	+.014
SQ REL. W/OTHERS	10.14	1.54	762	9.97	1.58	587	+1.96	+.054
SQ SELF	29.34	2.14	720	29.10	2.41	572	+1.85	+.051
SI READ	15.85	1.66	27	15.62	1.80	21	+0.46	+.069
SI ARITH	13.56	1.65	27	13.19	2.18	21	+0.66	+.097
SI TEACHER	8.63	1.47	27	8.14	1.42	21	+1.15	+.168
SI SCHOOL	9.04	1.08	26	8.15	1.90	20	+2.01	+.289
SI FUTURE SCHOOL	4.44	0.89	27	4.67	0.58	21	-0.99	-.145
SI REL. W/OTHERS	4.26	2.21	27	4.14	1.85	21	+0.19	+.028

TABLE III-17

STUDENT QUESTIONNAIRE AND INTERVIEW RESULTS
JACKSONVILLE, TIME 2

INDEX	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\bar{X}$	S.D.	N	$\bar{X}$	S.D.	N	t-val.	r_{pb}
SQ READ	15.16	1.60	754	14.67	1.74	693	+5.61	+.146
SQ ARITH	15.33	1.95	750	14.49	2.20	693	+7.66	+.198
SQ TEACHER	12.55	2.26	765	11.53	2.37	707	+8.45	+.198
SQ SCHOOL	3.77	0.51	773	3.56	0.64	710	+6.90	+.177
SQ FUTURE SCHOOL	9.49	0.97	755	9.57	0.88	705	-1.71	-.045
SQ REL. W/OTHERS	9.86	1.73	757	9.51	1.72	703	+3.93	+.102
SQ SELF	28.91	2.45	723	28.95	2.52	675	-0.33	-.010
SI READ	16.71	1.74	28	16.40	2.35	20	+0.53	+.078
SI ARITH	14.22	1.25	27	13.81	2.04	21	+0.86	+.126
SI TEACHER	8.48	1.40	27	8.48	1.81	21	+0.01	+.000
SI SCHOOL	9.36	1.03	28	8.67	1.93	21	+1.62	+.229
SI FUTURE SCHOOL	4.82	0.39	28	4.67	0.91	21	+0.81	+.117
SI REL. W/OTHERS	5.18	2.02	28	5.00	2.66	20	+0.26	+.039

students, at <u>both</u> times, than the CON school. The only exception to this is the "Orientation toward Future Schooling" index.

The puzzles for Jacksonville appear when the Student Interview results for the corresponding indexes are considered. For four of the six Interview-based indexes, there are no changes at all. For the remaining two, on the other hand, there is some evidence of change. The paradoxical aspect of this change is that it moves in the opposite direction from the change observed for the corresponding index based on the Student Questionnaire. For instance, the Questionnaire-based index for "Attitude toward Teacher" is higher for students in the EXP school at Time 1 ($t = 7.21$) and the gap becomes larger in favor of the EXP school at Time 2 ($t = 8.45$). However, when the corresponding index based on the Student Interview data is examined, it is found that the Time 1 results show a difference slightly in favor of the EXP school ($t = 1.15$), and that the Time 2 results show no difference at all ($t = 0.01$). A similar paradox occurs with regard to the "Orientation toward Future Schooling" index. The index based on the questionnaire data changes from a negligible positive t-value (+0.47) to a non-negligible negative t-value (-1.71); the corresponding index from the Interview results changes from marginal negative (-0.99) to marginal positive (+0.81). These results remain puzzling, and prevent the drawing of any clear conclusions as to the impacts of the incentives treatment upon student attitudes in Jacksonville.

Oakland is one of the sites at which the Parent-Teacher model was used. The results in Oakland exhibit an interesting pattern. These results are presented in Tables III-18 and III-19. The data from the Student Questionnaire indexes show several indications that the PT treatment is having a beneficial effect. The difference between EXP and CON on the Questionnaire-based indexes of (1) "Attitude toward Reading", (2) "Attitude toward Arithmetic", (3) "Attitude toward School", and (4) "Relations with Other Students" is in favor of the EXP school, and that the t-value for each of these indexes is larger at the second administration than at the first. This would suggest a positive impact of the PT

TABLE III-18

STUDENT QUESTIONNAIRE AND INTERVIEW RESULTS
OAKLAND, TIME 1

INDEX	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\bar{X}$	S.D.	N	$\bar{X}$	S.D.	N	t-val.	r_{pb}
SQ READ	15.12	1.70	389	14.66	1.86	529	+3.85	+.126
SQ ARITH	14.80	2.07	394	14.65	2.08	536	+1.10	+.036
SQ TEACHER	12.18	2.20	396	11.97	2.29	579	+1.41	+.045
SQ SCHOOL	3.68	0.58	399	3.67	0.60	577	+0.44	+.014
SQ FUTURE SCHOOL	9.70	0.66	398	9.70	0.70	573	+0.07	+.000
SQ REL. W/OTHERS	10.04	1.50	392	9.71	1.59	569	+3.22	+.103
SQ SELF	29.46	2.17	364	29.37	2.25	510	+0.57	+.020
SI READ	16.03	1.27	34	14.38	2.14	26	+3.72	+.439
SI ARITH	13.21	2.03	34	12.00	2.38	26	+2.12	+.268
SI TEACHER	8.14	1.88	35	8.14	1.49	26	-0.02	-.000
SI SCHOOL	8.86	1.40	35	7.69	1.95	26	+2.72	+.333
SI FUTURE SCHOOL	4.77	0.43	35	4.32	0.90	25	+2.59	+.322
SI REL. W/OTHERS	4.80	2.78	35	4.50	2.25	26	+0.45	+.058

TABLE III-19

STUDENT QUESTIONNAIRE AND INTERVIEW RESULTS
OAKLAND, TIME 2

INDEX	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
SQ READ	15.16	1.77	404	14.56	1.92	542	+4.93	+.158
SQ ARITH	15.16	1.99	400	14.77	2.14	544	+2.84	+.092
SQ TEACHER	11.71	2.43	410	11.61	2.36	545	+0.67	+.022
SQ SCHOOL	3.61	0.60	410	3.51	0.69	548	+2.39	+.077
SQ FUTURE SCHOOL	9.68	0.73	403	9.64	0.80	544	+0.88	+.028
SQ REL. W/OTHERS	10.00	1.53	411	9.53	1.60	551	+4.67	+.149
SQ SELF	29.49	2.04	379	29.29	2.48	508	+1.31	+.044
SI READ	16.04	1.07	23	14.84	1.97	25	+2.60	+.358
SI ARITH	14.00	1.41	23	13.62	1.70	26	+0.85	+.124
SI TEACHER	7.91	2.07	23	9.04	1.06	25	-2.41	-.335
SI SCHOOL	9.04	1.12	24	8.64	1.89	25	+0.90	+.130
SI FUTURE SCHOOL	4.67	0.92	24	4.81	0.40	26	-0.71	-.102
SI REL. W/OTHERS	5.00	2.60	25	4.40	1.87	25	+0.94	+.134

treatment on several student attitudes in Oakland, although some of the changes are not large. All of the changes, moreover, are in a direction opposite to any regression effect that might be operating.

When the Student Interview indexes are examined, however, the inference that there is a positive effect of the treatment is hardly supported. In the Student Interview results, there was only one of the six indexes for which even the direction of change was positive, and for that index the change is very small. Moreover, for four of the Interview-based indexes, the size of the negative change is substantial. This pattern, much like that for Jacksonville, precludes any clear inferences about the impact of the incentives model in Oakland, as far as student attitudes are concerned. Perhaps it should be noted that the shifts in direction on the Student Interview indexes are in the direction that would be predicted from a regression effect explanation, but this cannot really solve the puzzle.

One consistency between the two student attitude instruments in Oakland occurs on the "Attitude toward Teacher" index. The index based on the Questionnaire initially favors the EXP school (t-value at Time 1 is +1.41) and changes to being nearly equal for the two schools (t-value at Time 2 is +0.67). The corresponding index based on the Student Interview shows a difference of essentially zero at Time 1 (t-value is -0.02). At Time 2, however, there is a definite difference favoring the CON school (t-value is -2.41). This cannot be entirely explained by regression effect, since in the Interview data, the change is opposite to what a regression effect would predict.

The fourth of the individual site comparisons based on the Student Questionnaire and Student Interview data is that for San Antonio. These results are presented in Tables III-20 and III-21. This was, like Oakland, a site at which the PT model was used. Here, as with Oakland, the results seem erratic. They indicate, if anything, a generally negative impact of the treatment. In the Questionnaire-based index for "Attitude toward School", there is a change from a negligible difference at the first

TABLE III-20

STUDENT QUESTIONNAIRE AND INTERVIEW RESULTS
SAN ANTONIO, TIME 1

INDEX	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
SQ READ	14.95	1.54	395	15.07	1.54	288	-0.93	-.036
SQ ARITH	14.73	2.20	394	15.21	1.88	291	-2.98	-.113
SQ TEACHER	12.86	2.02	399	12.49	2.13	293	+2.32	+.088
SQ SCHOOL	3.77	0.47	400	3.74	0.52	292	+0.71	+.026
SQ FUTURE SCHOOL	9.66	0.76	397	9.69	0.70	289	-0.61	-.022
SQ REL. W/OTHERS	10.11	1.53	399	10.27	1.49	291	-1.43	-.055
SQ SELF	29.03	2.29	390	28.98	2.02	282	+0.30	+.010
SI READ	16.22	2.00	23	16.92	2.22	24	-1.13	-.166
SI ARITH	14.29	1.68	24	14.52	0.90	23	-0.58	-.087
SI TEACHER	8.83	1.63	24	8.75	1.80	24	+0.17	+.024
SI SCHOOL	8.71	2.37	24	9.79	0.83	24	-2.11	-.298
SI FUTURE SCHOOL	4.83	0.82	24	5.00	0.00	24	UNDEF	UNDEF
SI REL. W/OTHERS	4.46	2.47	24	4.63	2.37	24	-0.24	-.035

TABLE III-21

STUDENT QUESTIONNAIRE AND INTERVIEW RESULTS
SAN ANTONIO, TIME 2

INDEX	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
SQ READ	15.14	1.62	423	15.08	1.68	325	+0.46	+.017
SQ ARITH	15.08	2.04	418	15.49	1.70	324	-2.95	-.108
SQ TEACHER	12.70	1.88	419	12.26	2.10	328	+3.03	+.110
SQ SCHOOL	3.54	0.64	423	3.68	0.58	326	-3.11	-.113
SQ FUTURE SCHOOL	9.67	0.74	418	9.74	0.61	321	-1.56	-.057
SQ REL. W/OTHERS	9.72	1.65	420	10.07	1.58	328	-2.90	-.105
SQ SELF	28.93	2.24	411	29.01	2.19	320	-0.45	-.017
SI READ	16.27	2.47	22	17.91	1.57	22	-2.62	-.375
SI ARITH	14.43	1.08	23	15.00	0.00	23	-2.51	-.354
SI TEACHER	8.67	2.08	24	9.13	1.18	23	-0.94	-.138
SI SCHOOL	8.96	1.71	24	9.83	0.65	23	-2.28	-.322
SI FUTURE SCHOOL	5.00	0.00	24	5.00	0.00	23	UNDEF	UNDEF
SI REL. W/OTHERS	4.17	2.04	24	4.52	1.97	23	-0.61	-.090

administration between EXP and CON (t-value is +0.71), to a definite difference favoring the CON school at the second administration (t-value is -3.11). Another index from the Questionnaire which displays this same general pattern (a change from a slight difference at Time 1 to a difference favoring the CON school at Time 2) is the "Relations with Other Students" index. However, for this index, the change is small. The t-value at Time 1 is -1.43, and the t-value at Time 2 is -2.90.

The Student Interview data for San Antonio show a difference on the "Attitude toward School" index that favors the CON students, and the difference is stable over time. The t-value at Time 1 is -2.11, and the t-value at Time 2 is -2.28. The index for "Relations with Other Students" from the Student Interview shows no differences between the EXP and the CON schools at either time (the Time 1 t-value is -0.24, and the Time 2 t-value is -0.61). Of the remaining indexes, those for "Attitude toward Reading", "Attitude toward Arithmetic", and "Attitude toward Teachers", all show differences in favor of the CON group at Time 1 (or no differences) and those differences are increasingly in favor of the CON schools at Time 2 although in varying amounts. These changes are opposite in direction to a regression effect. Thus, the general picture for San Antonio, though fuzzy in detail, seems to indicate an impact on student attitudes opposite to that predicted.

Model-Based Comparisons--Student Questionnaire and Interview Results

In view of the generally inconsistent patterns obtained from the comparisons made at each of the individual sites, the interpretability of the model-based comparisons is lessened. However, the results of those comparisons are presented, with minimal discussion, since they are the best available summary information as to the overall impact of the different incentives models on the various student attitudes.

The results of the comparison for the pooled Teacher-Only sites (Cincinnati and Jacksonville) are presented in Tables III-22 and III-23. The Questionnaire-based index "Attitude toward School" changes in the

TABLE III-22

STUDENT QUESTIONNAIRE AND INTERVIEW RESULTS
TEACHER ONLY (TO), TIME 1

INDEX	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\bar{X}$	S.D.	N	$\bar{X}$	S.D.	N	t-val.	r_{pb}
SQ READ	14.97	1.71	1233	14.79	1.63	986	+2.44	+.052
SQ ARITH	15.26	1.77	1241	14.63	2.17	988	+7.53	+.157
SQ TEACHER	12.53	2.10	1251	12.14	2.21	996	+4.35	+.092
SQ SCHOOL	3.76	0.52	1261	3.71	0.55	997	+2.57	+.092
SQ FUTURE SCHOOL	9.64	0.73	1229	9.60	0.77	987	+0.04	+.022
SQ REL. W/OTHERS	10.15	1.51	1240	9.87	1.60	987	+4.15	+.088
SQ SELF	29.27	2.10	1192	29.12	2.34	970	+1.56	+.033
SI READ	15.78	1.60	51	15.98	1.83	42	-0.54	-.057
SI ARITH	13.63	1.87	51	13.74	1.91	42	-0.28	-.030
SI TEACHER	8.29	1.69	51	8.14	1.47	42	+0.45	+.048
SI SCHOOL	9.02	1.22	50	8.66	1.80	41	+1.14	+.120
SI FUTURE SCHOOL	4.69	0.71	51	4.81	0.45	42	-0.98	-.101
SI REL. W/OTHERS	4.20	2.15	51	4.21	2.10	42	-0.04	-.000

TABLE III-23

STUDENT QUESTIONNAIRE AND INTERVIEW RESULTS
TEACHER ONLY (TO), TIME 2

INDEX	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
SQ READ	14.88	1.76	1205	14.64	1.84	1084	+3.18	+.066
SQ ARITH	15.14	2.00	1201	14.53	2.23	1080	+6.91	+.143
SQ TEACHER	12.17	2.33	1223	11.67	2.40	1100	+5.08	+.105
SQ SCHOOL	3.63	0.62	1231	3.51	0.68	1104	+4.31	+.089
SQ FUTURE SCHOOL	9.55	0.88	1211	9.57	0.89	1092	-0.39	-.010
SQ REL. W/OTHERS	9.81	1.70	1210	9.51	1.68	1093	+4.31	+.089
SQ SELF	28.87	2.46	1156	28.88	2.48	1046	-0.09	-.000
SI READ	16.33	1.91	45	16.53	1.97	38	-0.45	-.050
SI ARITH	14.05	1.40	44	14.11	1.76	37	-0.18	-.020
SI TEACHER	8.27	1.64	45	8.51	1.62	39	-0.69	-.076
SI SCHOOL	9.00	1.35	46	8.85	1.69	39	+0.47	+.051
SI FUTURE SCHOOL	4.78	0.66	46	4.82	0.68	39	-0.26	-.028
SI REL. W/OTHERS	4.98	2.05	46	5.03	2.52	38	-0.10	-.010

predicted direction, and in a pattern not attributable to regression effect. The t-value at Time 1 is +2.57, and the t-value at Time 2 is 4.31. The few other changes that were observed are negative in direction, and potentially explainable as due to regression effect.

The model-based comparison for the pooled Parent-Teacher sites (Oakland and San Antonio) is presented in Tables III-24 and III-25. These results resemble the preceding ones in their lack of consistency. There are two positive changes observed in the Questionnaire-based indexes. These occur in "Attitude toward Reading" and in "Attitude toward Arithmetic". However, for these same two indexes on the interview, the direction of change is negative. This negative change might be due to a regression toward the mean, since in each case the shift observed is a movement toward decreasing differences. That is, for the Interview-based index of "Attitude toward Reading", the Time 1 t-value is +1.26 and the Time 2 t-value is -0.27. Similarly, for the Interview-based index of "Attitude toward Arithmetic", the Time 1 t-value is +1.17 and the Time 2 t-value is -0.17. The Interview-based index of "Attitude toward Teacher" shows a shift toward more favorable attitudes in the CON group. Here the Time 1 t-value is -0.05, and the Time 2 t-value is -2.30. Thus, this shift cannot be attributed to regression effect. However, the corresponding index from the Questionnaire shows a difference in favor of the EXP schools which is nearly identical at both times. The Time 1 t-value for the Questionnaire-based index of "Attitude toward Teacher" is +3.48, and the corresponding Time 2 value is +3.49. Thus, another inconsistency between the two instruments emerges.

The results from the comparison of the EXP schools in the PT sites with the EXP schools in the TO sites are presented in Tables III-26 and III-27. These tables reveal that there are a number of positive changes from the Questionnaire-based indexes, and two negative changes from the Interview-based indexes. The two negative changes on the Interview results (for "Attitude toward Reading" and for "Relations with Other Students") run directly counter to the changes on the corresponding Questionnaire-based indexes. However, for this comparison, any of the differences may

(Test resumes on page III-90)

TABLE III-24

STUDENT QUESTIONNAIRE AND INTERVIEW RESULTS
PARENT TEACHER (PT), TIME 1

INDEX	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\bar{X}$	S.D.	N	$\bar{X}$	S.D.	N	t-val.	r_{pb}
SQ READ	15.03	1.71	784	14.80	1.77	817	+2.65	+.066
SQ ARITH	14.77	2.14	788	14.85	2.03	827	-0.78	-.020
SQ TEACHER	12.52	2.14	795	12.14	2.25	872	+3.48	+.085
SQ SCHOOL	3.73	0.53	799	3.69	0.58	869	+1.27	+.032
SQ FUTURE SCHOOL	9.68	0.71	795	9.70	0.70	862	-0.48	-.010
SQ REL. W/OTHERS	10.07	1.51	791	9.90	1.58	860	+2.26	+.056
SQ SELF	29.23	2.24	754	29.23	2.18	792	+0.04	+.000
SI READ	16.11	1.59	57	15.60	2.51	50	+1.26	+.122
SI ARITH	13.66	1.95	58	13.18	2.22	49	+1.17	+.113
SI TEACHER	8.42	1.80	59	8.44	1.66	50	-0.05	-.000
SI SCHOOL	8.80	1.84	59	8.70	1.84	50	+0.27	+.026
SI FUTURE SCHOOL	4.80	0.61	59	4.65	0.72	49	+1.12	+.108
SI REL. W/OTHERS	4.66	2.64	59	4.56	2.29	50	+0.21	+.020

TABLE III-25

STUDENT QUESTIONNAIRE AND INTERVIEW RESULTS
PARENT TEACHER (PT), TIME 2

INDEX	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
SQ READ	15.15	1.69	827	14.76	1.85	867	+4.57	+.110
SQ ARITH	15.11	2.01	818	15.04	2.01	868	+0.80	+.020
SQ TEACHER	12.21	2.23	829	11.85	2.29	873	+3.29	+.079
SQ SCHOOL	3.57	0.62	833	3.57	0.65	874	+0.05	+.000
SQ FUTURE SCHOOL	9.67	0.73	821	9.68	0.74	865	-0.11	-.000
SQ REL. W/OTHERS	9.86	1.60	831	9.73	1.62	879	+1.72	+.041
SQ SELF	29.20	2.16	790	29.18	2.38	828	+0.20	+.000
SI READ	16.16	1.87	45	16.28	2.36	47	-0.27	-.028
SI ARITH	14.22	1.26	46	14.27	1.41	49	-0.17	-.017
SI TEACHER	8.30	2.08	47	9.08	1.11	48	-2.30	-.232
SI SCHOOL	9.00	1.43	48	9.21	1.54	48	-0.69	-.071
SI FUTURE SCHOOL	4.83	0.66	48	4.90	0.31	49	-0.62	-.063
SI REL. W/OTHERS	4.59	2.35	49	4.46	1.90	48	+0.31	+.032

TABLE III-26

STUDENT QUESTIONNAIRE AND INTERVIEW RESULTS
PARENT TEACHER EXPERIMENTAL (PT-E) AND TEACHER ONLY EXPERIMENTAL (TO-E), TIME 1

INDEX	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
SQ READ	15.03	1.71	784	14.97	1.71	1233	+0.87	+.020
SQ ARITH	14.77	2.14	788	15.26	1.77	1241	-5.66	-.125
SQ TEACHER	12.52	2.14	795	12.53	2.10	1251	-0.17	-.000
SQ SCHOOL	3.73	0.53	799	3.76	0.52	1261	-1.57	-.035
SQ FUTURE SCHOOL	9.68	0.71	795	9.64	0.73	1229	+1.28	+.028
SQ REL. W/OTHERS	10.07	1.51	791	10.15	1.51	1240	-1.03	-.022
SQ SELF	29.23	2.24	754	29.27	2.10	1192	-0.37	-.010
SI READ	16.11	1.59	57	15.78	1.60	51	+1.04	+.101
SI ARITH	13.66	1.95	58	13.63	1.87	51	+0.01	+.010
SI TEACHER	8.42	1.80	59	8.29	1.69	51	+0.39	+.037
SI SCHOOL	8.80	1.84	59	9.02	1.22	50	-0.73	-.071
SI FUTURE SCHOOL	4.80	0.61	59	4.69	0.71	51	+0.88	+.084
SI REL. W/OTHERS	4.66	2.64	59	4.70	2.15	51	+1.00	+.096

TABLE III-27

STUDENT QUESTIONNAIRE AND INTERVIEW RESULTS
PARENT TEACHER EXPERIMENTAL (PT-E) AND TEACHER ONLY EXPERIMENTAL (TO-E), TIME 2

INDEX	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
SQ READ	15.15	1.69	827	14.88	1.76	1205	+3.49	+.077
SQ ARITH	15.11	2.01	818	15.14	2.00	1201	-0.31	-.000
SQ TEACHER	12.21	2.23	829	12.17	2.33	1223	+0.38	+.010
SQ SCHOOL	3.57	0.62	833	3.63	0.62	1231	-2.02	-.045
SQ FUTURE SCHOOL	9.67	0.73	821	9.55	0.88	1211	+3.28	+.073
SQ REL. W/OTHERS	9.86	1.60	831	9.81	1.70	1210	+0.66	+.014
SQ SELF	29.20	2.16	790	28.87	2.46	1156	+3.06	+.069
SI READ	16.16	1.87	45	16.33	1.91	45	-0.45	-.048
SI ARITH	14.22	1.26	46	14.05	1.40	44	+0.61	+.066
SI TEACHER	8.30	2.08	47	8.27	1.64	45	+0.08	+.010
SI SCHOOL	9.00	1.43	48	9.00	1.35	46	+0.00	+.000
SI FUTURE SCHOOL	4.83	0.66	48	4.78	0.66	46	+0.37	+.039
SI REL. W/OTHERS	4.59	2.35	49	4.98	2.05	46	-0.85	-.088

well be due to differences in historical events at the different sites, and so should not be definitely associated with the differential impact of the Parent-Teacher model.

The final set of comparisons based on the Student Questionnaire and Student Interview data is the comparison in which all the EXP schools are pooled and compared with the pooled CON schools. This comparison suppresses any possible differences between sites or differences between the two incentives models. The results from this overall comparison are presented in Tables III-28 and III-29. They indicate some positive change in the Questionnaire-based index of "Attitude toward Reading", and some negative change in the Questionnaire-based index of "Orientation toward Self". In addition, this overall comparison indicates that there are several negative changes in the Interview-based indexes. The indexes for "Attitude toward School", "Attitude toward Reading", "Attitude toward Arithmetic", and "Attitude toward Teachers" all show changes in the direction opposite to prediction. Some, but not all, of these changes can be explained by possible regression effects.

Discussion--Student Questionnaire and Interview Results

Perhaps the most interesting outcome of this analysis is not the information which emerges as to the particular pattern of student attitudes and attitude change at the different sites and under the different treatments, but instead the high degree of inconsistency between the data collected from the Student Questionnaire instrument and the Student Interview instrument. This lack of consistency makes it impossible to draw substantive conclusions from this data with any confidence, but does indicate two obvious points of caution that must be noted in any effort to carry out further studies of this general type. The first point is technical, and simply stated, is the following: The results of attitude surveys are dependent, to a fairly high degree, upon the instrument and the circumstances of its administration. Thus, if clear and precise comparisons are desired, it will be necessary to invest considerable energy

TABLE III-28

STUDENT QUESTIONNAIRE AND INTERVIEW RESULTS
EXPERIMENTAL AND CONTROL (E/C) ALL, TIME 1

INDEX	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\bar{X}$	S.D.	N	$\bar{X}$	S.D.	N	t-val.	r_{pb}
SQ READ	14.99	1.71	2017	14.80	1.69	1803	+3.54	+.057
SQ ARITH	15.07	1.94	2029	14.73	2.11	1815	+5.18	+.083
SQ TEACHER	12.53	2.12	2046	12.14	2.23	1868	+5.59	+.089
SQ SCHOOL	3.75	0.53	2060	3.70	0.56	1866	+2.89	+.046
SQ FUTURE SCHOOL	9.65	0.72	2024	9.65	0.74	1849	+0.29	+.000
SQ REL. W/OTHERS	10.12	1.51	2031	9.89	1.59	1847	+4.66	+.075
SQ SELF	29.26	2.16	1946	29.17	2.27	1762	+1.19	+.020
SI READ	15.95	1.60	108	15.77	2.22	92	+0.67	+.048
SI ARITH	13.64	1.90	109	13.44	2.09	91	+0.72	+.051
SI TEACHER	8.36	1.74	110	8.30	1.57	92	+0.25	+.017
SI SCHOOL	8.90	1.58	109	8.68	1.31	91	+0.91	+.064
SI FUTURE SCHOOL	4.75	0.66	110	4.73	0.62	91	+0.22	+.017
SI REL. W/OTHERS	4.45	2.43	110	4.40	2.20	92	+0.13	+.010

TABLE III-29

STUDENT QUESTIONNAIRE AND INTERVIEW RESULTS
EXPERIMENTAL AND CONTROL (E/C) ALL, TIME 2

INDEX	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
SQ READ	14.99	1.73	2032	14.69	1.84	1951	+5.25	+.083
SQ ARITH	15.13	2.00	2019	14.76	2.15	1948	+5.68	+.090
SQ TEACHER	12.19	2.29	2052	11.75	2.35	1973	+5.96	+.093
SQ SCHOOL	3.61	0.62	2064	3.54	0.67	1978	+3.33	+.052
SQ FUTURE SCHOOL	9.60	0.82	2032	9.62	0.83	1957	-0.55	-.010
SQ REL. W/OTHERS	9.83	1.66	2041	9.61	1.66	1972	+4.33	+.068
SQ SELF	29.00	2.35	1946	29.01	2.44	1874	-0.09	-.000
SI READ	16.24	1.88	90	16.39	2.18	85	-0.47	-.036
SI ARITH	14.13	1.33	90	14.20	1.56	86	-0.30	-.022
SI TEACHER	8.28	1.87	92	8.83	1.38	87	-2.21	-.164
SI SCHOOL	9.00	1.38	94	9.05	1.61	87	-0.21	-.014
SI FUTURE SCHOOL	4.81	0.66	94	4.86	0.51	88	-0.63	-.047
SI REL. W/OTHERS	4.78	2.21	95	4.71	2.20	86	+0.21	+.017

in developing the survey instruments, ascertaining their sensitivity to random influences, and then minimizing these influences in the actual data collection. The second point is procedural: Field projects of this type require large amounts of lead-time for design and the exercise of authority to ensure that the original design is followed rigorously.

To make this point somewhat more specific, we can simply list some of the sources of possible inconsistency between the results from the Student Questionnaire and the Student Interview. These sources include:

1. Differences in the content of individual questions
2. Differences in the format by which corresponding questions are presented
3. Differences in the administrative circumstances in which the data is collected (The most obvious such difference in this data is that of the time difference between the two administrations. Other differences include differences in the research worker's personality, the physical surroundings, *etc.*)
4. An error in the data processing, at any of the several stages involved.

In addition to these technical possibilities for introducing distortion into the real pattern, there are two important and fairly plausible substantive possibilities. These are:

1. The students' opinion and attitude on these topics may in fact change rapidly, depending mainly upon recent events.
2. The respondent (student) really does not have a response that truly describes him on these topics. They simply are not matters on which he has developed an orientation, yet he feels constrained to provide an answer.

Since the respondents here are fairly young students, these two possibilities cannot be dismissed without investigation. However, such investigation has not been carried out by previous workers, and was not feasible with the data and time constraints of the present project. Thus, these issues remain unresolved.

Student Behavioral Results--Introduction

The student behavioral data consists of material obtained from school documents, namely the average classroom attendance rates, and material collected by research workers during structured classroom observations. The attendance data will be discussed first.

The classroom attendance data collected was coded for monthly summaries for each full month from October, 1971, through April, 1972, the full months during which the incentive project was in operation. For each month, in each classroom, three variables were created:

1. Total number of student days of absence per month
2. Average number of students enrolled for the month
3. Number of instructional days in the month.

These data were put into the following formula to calculate an index for each month:

$$\text{Absenteeism index} = \frac{\text{total number of absences observed}}{\text{maximum possible number of absences}}$$

$$= \frac{\text{(variable 1)}}{\text{(variable 2) (variable 3)}}$$

These indexes then were used in the analysis. It should be emphasized that the unit here is the classroom, not the student. Thus, a figure of 7.36 in the tables means that the absence rate was 7.36% for that set of classrooms in that month. It should also be noted that the direction of differences in the tables is numerically consistent with tables in the other sections. However, since <u>low</u> absence rates are more favorable than high ones, it is a <u>negative</u> t-value that indicates a difference in favor of the EXP school. Finally, summary indices of absence during the fall term were constructed simply by adding together the absence rate for November and December. Similar summary absence rates for the spring term were constructed by averaging together the absence rates for March and April. The results of the analyses of attendence are first presented separately for each of the four individual sites.

Student Attendance Results--Single-Site Comparison

The results of the attendance analysis for Cincinnati are presented in Table III-30. They indicate that in the fall term, the absenteeism at the EXP school was lower than at the CON school, but that in the spring, this difference had disappeared. In fact, in the spring, the absenteeism in the EXP school was higher than that of the CON school. This, then, seems to be an effect, though small in magnitude, in the direction opposite to that predicted by the hypothesis. However, this trend is in accord with the trend of the achievement data in Cincinnati. Inspection of the month-by-month results provides no grounds to qualify the conclusion further.

TABLE III-30

ABSENTEEISM DATA--CINCINNATI

MONTH(S)	ABSENTEEISM (%)		DIFF (%) (EXP-CON)	t-val.	DIFF IN FAVOR
	EXP	CON			
November	6.86	7.21	-0.35	-0.34	EXP
December	6.94	9.99	-3.05	-2.67	EXP
January	10.69	9.48	+1.21	+0.84	CON
February	9.17	9.52	-0.35	-0.28	EXP
March	9.04	8.79	+0.25	+0.18	CON
April	9.81	7.48	+2.33	+1.67	CON
Fall Term	6.90	8.60	-1.70	-1.81	EXP
Spring Term	9.43	8.14	+1.29	+0.99	CON

Next, the results of the analysis of attendance patterns in Jacksonville are examined. These results are given in Table III-31. They indicate no consistent impact of the treatment on attendance. That is, in the fall term, the rate of absenteeism at the EXP school is slightly lower than the rate at the CON school. In the spring, this same direction of difference exists, but the gap is very slightly smaller. Thus, the direction of change is, as in Cincinnati, opposite to that predicted by the hypothesis. The size of the change, however, is negligible. Inspection of the month-by-month data indicates an erratic pattern. On balance, the conclusion is that there is no effect of the treatment on attendance rates in Jacksonville.

TABLE III-31

ABSENTEEISM DATA--JACKSONVILLE

MONTH(S)	ABSENTEEISM (%)		DIFF (%)		DIFF IN
	EXP	CON	(EXP-CON)	t-val.	FAVOR
November	4.67	5.69	-1.02	-1.87	EXP
December	6.80	7.75	-0.95	-1.16	EXP
January	8.18	7.76	+0.42	+0.66	CON
February	7.75	8.61	-0.86	-1.12	EXP
March	5.43	5.92	-0.49	-0.68	EXP
April	4.84	5.73	-0.89	-1.47	EXP
Fall Term	5.74	6.72	-0.98	-1.67	EXP
Spring Term	5.14	5.83	-0.69	-1.13	EXP

The results for Oakland are presented in Table III-32. In Oakland, the pattern is not entirely clear, but the general trend is in the direction predicted by the hypotheses. At both the fall term and the spring term observations, the absence rate at the EXP school is higher than that at the CON school. However, the gap between the two schools is smaller in the spring than in the fall. This improvement in the level of absenteeism at the EXP school relative to that of the CON school is in the direction predicted, even though the absenteeism at the EXP school remains slightly higher than that at the CON school. The impression gained from comparing the fall term index to the spring term index is confirmed by an examination of the month-by-month data, although this data also reveal that most of the effect is due to an unusually low absence rate at the EXP school during the month of March.

TABLE III-32

ABSENTEEISM DATA--OAKLAND

MONTH(S)	ABSENTEEISM (%) EXP	ABSENTEEISM (%) CON	DIFF (%) (EXP-CON)	t-val.	DIFF IN FAVOR
November	7.39	5.32	+2.07	+2.74	CON
December	8.44	7.29	+1.15	+1.11	CON
January	10.38	7.39	+2.99	+3.06	CON
February	8.76	6.41	+2.35	+2.59	CON
March	5.19	6.36	-1.17	-2.11	EXP
April	10.32	8.19	+2.13	+1.89	CON
Fall Term	7.92	6.31	-1.61	+2.12	CON
Spring Term	7.76	7.28	-0.48	+0.67	CON

The attendance patterns in San Antonio are summarized in Table III-33. The results from the fall term index and spring term index show a nearly random pattern but one that is opposite to the direction predicted by the hypothesis. That is, in the fall, absenteeism at the EXP school is very slightly lower than it is at the CON school. However, in the spring, absenteeism at the EXP school is very slightly higher than it is at the CON school. These differences are so small as to be negligible, but they are in the direction contrary to the hypothesis. A study of the detailed data tends to confirm the impression that the differences are essentially random in this case. Thus, the conclusion that there was no treatment impact on attendance in San Antonio.

TABLE III-33

ABSENTEEISM DATA--SAN ANTONIO

MONTH(S)	ABSENTEEISM (%)		DIFF (%)		DIFF IN
	EXP	CON	(EXP-CON)	t-val.	FAVOR
November	6.23	5.98	+0.25	+0.23	CON
December	7.11	7.69	-0.58	-2.67	EXP
January	10.68	10.62	+0.06	+0.04	CON
February	8.38	7.68	+0.07	+0.52	CON
March	7.83	6.26	+1.57	+1.42	CON
April	6.09	7.23	-1.14	-0.86	EXP
Fall Term	6.67	6.83	-0.16	-0.16	EXP
Spring Term	6.96	6.75	+0.21	+0.19	CON

In general, the comparison of attendance rates at the four individual sites reveals a pattern of differences between EXP and CON schools that probably is best regarded as random. In the two sites which received the Teacher-Only model, Cincinnati and Jacksonville, there are no differences large enough to mention. In Oakland, there is a small difference in the predicted direction. However, in San Antonio (which along with Oakland received the Parent-Teacher model), there is a small difference opposite to the predicted direction. The general conclusion that must be drawn is that the application of the incentives models had no discernible impact on student attendance.

Student Attendance Results--Model-Based Comparisons

Because of the essentially nil evidence of any impact of the treatment on attendance rates in the individual sites, no attempt will be made to discuss or interpret the results obtained from the model-based comparisons of attendance rates. Such interpretations would be unwarranted, in light of the erratic individual site patterns. However, the results of such comparisons have been calculated, and these results are presented in Table III-34, for the Teacher-Only sites combined; Table III-35, for the Parent-Teacher sites combined; Table III-36, for the EXP schools in the PT sites versus the EXP schools in the TO sites; and, finally, Table III-37, for the overall comparison of all EXP schools versus all CON schools.

Student Attendance Results--Discussion

There are several possible explanations for the general lack of impact of the incentives treatments on student attendance rates. Two explanations are substantive and one is procedural. One substantive possibility is that the introduction of the incentives treatment *per se* did not encourage a student to attend more regularly during the year. School staff members might knowingly ignore students who were chronic absentees, because, in their absence, there would be more time available for working

TABLE III-34

ABSENTEEISM DATA--E/C TEACHER ONLY

MONTH(S)	ABSENTEEISM (%)		DIFF (%)	t-val.	DIFF IN FAVOR
	EXP	CON	(EXP-CON)		
November	5.50	6.34	-0.84	-1.53	EXP
December	6.86	8.71	-1.85	-2.73	EXP
January	9.13	8.50	+0.63	+0.88	CON
February	8.28	9.00	-0.72	-1.05	EXP
March	6.69	7.15	-0.46	-0.60	EXP
April	6.57	6.48	+0.09	+0.12	CON
Fall Term	6.18	7.52	-1.34	-2.52	EXP
Spring Term	6.63	6.81	-0.18	-0.26	EXP

TABLE III-35

ABSENTEEISM DATA--E/C PARENT TEACHER

MONTH(S)	ABSENTEEISM (%)		DIFF (%)	t.val.	DIFF IN FAVOR
	EXP	CON	(EXP-CON)		
November	6.83	5.55	+1.28	+2.02	CON
December	7.80	7.43	+3.70	+0.47	CON
January	10.53	8.53	+2.00	+2.31	CON
February	8.58	6.85	+1.73	+2.29	CON
March	6.47	6.33	+0.14	+0.24	CON
April	8.27	7.86	+0.41	+0.45	CON
Fall Term	7.31	6.50	+0.81	+1.33	CON
Spring Term	7.37	7.10	+0.27	+0.45	CON

TABLE III-36

ABSENTEEISM DATA--PARENT TEACHER-E/TEACHER ONLY-E

MONTH(S)	ABSENTEEISM (%)		DIFF (%)	t-val.	DIFF IN FAVOR
	EXP	CON	(EXP-CON)		
November	6.83	5.50	+1.33	+2.14	CON
December	7.80	6.86	+0.94	+1.26	CON
January	10.53	9.13	+0.14	+1.58	CON
February	8.58	8.28	+0.30	+0.40	CON
March	6.47	6.69	-0.22	-0.32	EXP
April	8.27	6.57	+1.70	+1.88	CON
Fall Term	7.31	6.18	+1.13	+1.87	CON
Spring Term	7.37	6.63	+0.74	+1.05	CON

TABLE III-37

ABSENTEEISM DATA--E/C ALL SITES

MONTH(S)	ABSENTEEISM (%)		DIFF (%)	t-val.	DIFF IN FAVOR
	EXP	CON	(EXP-CON)		
November	6.06	5.97	+0.09	+0.22	CON
December	7.26	8.11	-0.85	-1.66	EXP
January	9.72	8.51	+1.21	+2.17	CON
February	8.41	7.98	-0.43	+0.83	CON
March	6.59	6.76	-0.17	-0.33	EXP
April	7.31	7.14	+0.17	+0.29	CON
Fall Term	6.66	7.04	-0.38	-0.93	EXP
Spring Term	6.95	6.95	+0.00	+0.00	---

with the students who are present. This, then, is one possible explanation as to why there would be no effort prompted by the incentive treatment to influence the attendance of a student who is chronically absent. On the other hand, for the student who is absent only occasionally, the pattern and possible causes are unpredictable, and hence there probably is little that the school staff can do to improve his attendance.

This brings up the second substantive explanation. The percentages of absence indicated in this data are considerably lower than those rates thought to exist in similar low-income schools. This could itself indicate a variety of processes at work. Both the EXP and CON schools have very good attendance throughout the year and the small amount of absenteeism that does occur in these schools is probably due to real illness and other causes which have nothing to do with school-induced motivation or other pressure. So, there would be a ceiling effect operative on attendance. No treatment could improve the attendance much in view of its already high level.

The third explanation is that the CON school faculties and principals worked harder at reducing absenteeism than they would have normally. They were not naive subjects in this project; and the impact of their control status, daily reinforced by the TAC and MAC presence, might have been sufficient to focus their attention on reducing absenteeism. This factor, then, would make the CON absenteeism rate a difficult benchmark against which to measure the net impact of offering incentives in the EXP school.

Student Classroom Observation Results--Introduction

The data to be discussed in this section were obtained from structured classroom observation, conducted by workers trained and supervised by TAC. The observations took place in December (Time 1) and May (Time 2). At each of these time periods, an observer spent a total of approximately 80 minutes observing each class. The observations were made in four sessions of about 20 minutes each, in order to avoid any distortion that might be caused as a result of observing one atypical class session. The

data were coded by means of a system devised for this project. The system included observations of behaviors of the teacher, behaviors of several students, and the focus of class activity.

Briefly, the procedure used to record observation data was as follows. The observer began by choosing one student at random as a subject to be observed. For that student, the observer coded the kind of activity the student was engaged in during a ten-second period and the context of the activity. There were 14 different activity categories allowable and four context categories. These are presented in Table III-38. When ten seconds had elapsed, the observer chose another student at random, and recorded that student's behavior for a ten-second interval. The observer continued in this fashion for three minutes, thus sampling the behavior of a number of students. After three minutes, the observer turned his attention to the teacher. Each ten seconds, the observer recorded the activity of the teacher (see Table III-38 for the possible codes for teacher activity), and the focus of that activity. Then, after three minutes of recording the teacher's behavior and its context, the observer again switched to recording the student behaviors. The observation of a single session thus gave approximately six minutes of observation of the students, based on a number of students, and approximately six minutes of observation of the teacher, with approximately five minutes for getting set up and leaving.

The data were recorded simply by counting the frequency of each activity and context separately for the teacher and student observations. For the analysis, the separate sessions within a single time point were combined, and these frequencies are the data presented in the tables that follow. Thus, the unit of observation is the classroom. However, in calculating the results shown in the tables here, no distinction was made as to differences between teacher, or differences within grade. That is, all tables show totals by school or sets of schools.

TABLE III-38

CLASSROOM OBSERVATION ACTIVITIES AND CONTEXTS: STUDENT

ITEM	ACTIVITY	EXPLANATION	HYPO-THESES
1	Reading		E > C
2	Referring to Reading	Student makes reference to personal reading.	E > C
3	Writing		E > C
4	Reciting	Student is delivering prepared and/or rote material.	-----
5	Listening		E < C
6	Being Tested		E > C
7	Asking Academic Questions		E > C
8	Answering Academic Questions		-----
9	Criticizing		E < C
10	Complimenting		E > C
11	Initiating	Student starts academic work on his own	E > C
12	Academic Game Playing	Spelling bees, etc.	E > C
13	Nonproductive	Goofing off, disrupting, etc.	E < C
14	Productive	Behavior is not related to class but not disruptive.	E > C
Context			
1	Alone		E > C
2	Small Group	2-5 students	E > C
3	Large Group	6 students; 3/4 class	E < C
4	Whole Class		E < C

Single-Site Comparisons--Student Classroom Observation Results

Observation Results. The data from each site are first presented in Tables III-39 through III-46. A formal test of significance has not been employed in the comparisons that follow, but rather patterns of change were searched for thorough comparisons of the t-value at Time 1 with the corresponding t-value at Time 2.

Cincinnati results are found in Tables III-39 and III-40, the Time 1 and Time 2 observations, respectively, In Cincinnati, only Item 13 "Nonproductive", shows a clear change, with there being markedly more of that student classroom behavior in the EXP school as compared to the CON school toward the end of the year (t-value = +3.24). In this case, the EXP scores remained about the same, whereas there was a significant decrease in CON school nonproductive behavior from Time 1 to Time 2. This change in status in the area of classroom management was also observed more generally by MAC. They reported that the EXP school's principal had been ill for extended periods during the year which had probably caused a deterioration in school discipline. On the other hand, the principal of the CON school had regarded the project competitively; perhaps this decline in nonproductive behavior was one consequence of a general tightening up in the CON school. This trend in Item 13 is somewhat confirmed by the decline of the EXP school with a concomitant increase of the CON school in Item 14, "Other Productive". The frequency of this behavior in the EXP school was reduced by one-half whereas it doubled in the CON school. Both of these trends are contrary to what was hypothesized.

On the other hand, changes in Item 1, "Reading"; Item 6, "Being Tested"; and Item 7, "Asking Questions" were in the direction hypothesized. These three changes are very slight, involving only t-value changes from negative to positive from Time 1 to Time 2 with no significant difference in t-values in favor of the EXP school at Time 2. Only in "Being Tested" is there any appreciable increase in the frequency in the EXP school from Time 1 to Time 2. The t-value changes in "Reading" and "Asking Questions" are due to decreases in these activities in the CON school.

TABLE III-39

STUDENT CLASSROOM OBSERVATION

E/C, CINCINNATI, TIME 1

ITEMS & CONTEXT	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\bar{X}$	S.D.	N	$\bar{X}$	S.D.	N	t-val.	r_{pb}
ITEM 1	4.11	5.24	56	5.23	5.86	64	-1.10	-.101
ITEM 2	0.00	0.00	56	0.08	0.63	64	UNDEF	UNDEF
ITEM 3	10.20	8.00	56	9.53	6.86	64	+0.49	+.045
ITEM 4	2.80	5.23	56	2.86	4.29	64	-0.06	-.000
ITEM 5	9.16	7.65	56	7.80	6.50	64	+1.06	+.097
ITEM 6	0.32	2.40	56	0.97	5.37	64	-0.83	-.076
ITEM 7	0.71	1.46	56	1.19	3.15	64	-1.03	-.094
ITEM 8	3.61	4.49	56	3.58	3.90	64	+0.04	+.000
ITEM 9	0.04	0.19	56	0.03	0.18	64	+0.13	+.014
ITEM 10	0.00	0.00	56	0.00	0.00	64	UNDEF	UNDEF
ITEM 11	0.05	0.22	56	0.16	0.91	64	-0.82	-.075
ITEM 12	0.32	1.03	56	0.42	1.85	64	-0.36	-.033
ITEM 13	3.27	3.51	56	2.28	4.59	64	+1.31	+.120
ITEM 14	1.12	2.69	56	0.20	0.90	64	+2.61	+.234
ONE STUD	2.73	6.16	56	4.42	7.80	64	-1.30	-.011
SM GROUP	4.05	6.11	56	2.33	4.71	64	+1.74	+.016
LG GROUP	6.39	9.50	56	12.56	13.74	64	-2.82	-.251
WHOL CLAS	22.88	13.76	56	15.22	15.32	64	+2.86	+.255

TABLE III-40

STUDENT CLASSROOM OBSERVATION
E/C, CINCINNATI, TIME 2

ITEMS & CONTEXT	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
ITEM 1	4.68	6.72	72	3.79	6.04	80	+0.86	+.070
ITEM 2	0.21	1.77	72	0.05	0.35	80	+0.78	+.064
ITEM 3	7.49	7.72	72	5.95	6.62	80	+1.32	+.107
ITEM 4	1.11	2.69	72	1.03	2.32	80	+0.21	+.017
ITEM 5	3.97	5.07	72	5.08	6.23	80	-1.19	-.096
ITEM 6	1.04	4.74	72	0.60	2.98	80	+0.69	+.060
ITEM 7	0.87	1.85	72	0.68	1.48	80	+0.74	+.057
ITEM 8	2.90	4.46	72	3.35	4.88	80	-0.59	-.048
ITEM 9	0.31	1.04	72	0.16	0.65	80	+1.03	+.084
ITEM 10	0.08	0.50	72	0.00	0.00	80	UNDEF	UNDEF
ITEM 11	0.00	0.00	72	0.04	0.34	80	UNDEF	UNDEF
ITEM 12	0.21	0.79	72	0.25	1.91	80	-0.17	-.014
ITEM 13	3.14	4.56	72	1.23	2.54	80	+3.24	+.256
ITEM 14	0.65	1.58	72	0.40	1.35	80	+1.07	+.087
ONE STUD	7.56	8.38	72	4.60	6.77	80	+2.40	+.193
SM GROUP	3.06	6.20	72	1.99	4.58	80	+1.22	+.099
LG GROUP	5.71	9.89	72	9.70	12.25	80	-2.19	-.176
WHOL CLAS	10.01	14.02	72	6.19	11.05	80	+1.88	+.152

Finally, the Cincinnati data indicate that there is a rather clear trend toward individual student attention in the EXP school classrooms. In the context, "One Student", the t-value went from -1.30 at Time 1 to +2.40 at Time 2. While the differences in the frequency of small and large group instruction remains about the same from Time 1 to Time 2, there is a relatively greater decline of "Whole Class" instruction in the EXP school.

Tables III-41 and III-42 present the Time 1 and Time 2 Student Classroom Observation data for Jacksonville. There were relatively greater decreases of student behaviors in the EXP school in Item 7, "Asking Questions", and in Item 14,"Productive". Both of these changes were not in the direction hypothesized. In five items there were relative changes in the direction hypothesized: Item 1, "Reading"; Item 3, "Writing"; Item 8, "Answering Questions"; Item 9, "Criticizing"; and Item 12, "Academic Game Playing". In "Reading" the change is most pronounced with the t-values changing from -1.83 at Time 1 to +3.55 at Time 2. As can be seen from the means for Item 1, the frequency of this activity nearly doubled in the EXP school while it was being reduced in the CON school. The relative changes in these other four items are due either to marginal increases in the EXP school or decreases in the CON school and do not reveal any clear pattern.

The contexts of instruction in Jacksonville changed opposite to the direction hypothesized in two areas, "One Student" and "Large Group". In the "One Student" context, the frequency of this mode of instruction in the EXP school was reduced by one-half from Time 1 to Time 2 (t-values of +2.57 and -0.96, respectively). At the same time, "Large Group" instruction doubled in the EXP school from a Time 1 mean of 5.40 to a Time 2 mean of 11.08. Although there was an increase in "Large Group" instruction in the CON school, it did not approach the magnitude of the EXP school increase. Finally, it is worth noting that the frequency of "Whole Class" instruction increased somewhat in the EXP school; and although the t-values are small, they did change in the direction opposite to what was hypothesized. Therefore, the pattern in the Jacksonville EXP school

TABLE III-41

STUDENT CLASSROOM OBSERVATION
E/C, JACKSONVILLE, TIME 1

ITEMS & CONTEXT	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\bar{X}$	S.D.	N	$\bar{X}$	S.D.	N	t-val.	r_{pb}
ITEM 1	3.79	4.82	110	5.22	5.85	79	-1.83	-.133
ITEM 2	0.35	1.08	110	0.00	0.00	79	UNDEF	UNDEF
ITEM 3	7.27	6.52	110	7.39	6.17	79	-0.13	-.010
ITEM 4	3.12	3.78	110	3.24	3.77	79	-0.22	-.017
ITEM 5	6.92	6.38	110	7.13	5.87	79	-0.23	-.017
ITEM 6	1.02	4.05	110	0.86	3.19	79	+0.29	+.020
ITEM 7	1.50	2.14	110	1.16	2.02	79	+1.09	+.077
ITEM 8	3.26	3.61	110	3.71	3.84	79	-0.81	-.059
ITEM 9	0.64	1.81	110	0.35	0.95	79	+1.26	+.092
ITEM 10	0.08	0.43	110	0.00	0.00	79	UNDEF	UNDEF
ITEM 11	0.39	0.89	110	0.22	0.71	79	+1.45	+.106
ITEM 12	1.39	3.47	110	2.29	6.33	79	-1.25	-.091
ITEM 13	4.77	5.02	110	4.20	4.89	79	+0.78	+.057
ITEM 14	1.20	3.80	110	0.24	1.19	79	+2.38	+.171
ONE STUD	12.34	8.92	110	9.18	7.40	79	+2.58	+.185
SM GROUP	4.44	6.76	110	4.66	7.18	79	-0.22	-.017
LG GROUP	5.41	8.75	110	8.13	10.94	79	-1.89	-.137
WHOL CLAS	13.82	12.26	110	14.04	13.55	79	-0.12	-.010

TABLE III-42

STUDENT CLASSROOM OBSERVATION
E/C, JACKSONVILLE, TIME 2

ITEMS & CONTEXT	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
ITEM 1	6.43	5.42	112	3.95	4.65	100	+3.55	+.238
ITEM 2	0.01	0.09	112	0.01	0.10	100	-0.08	-.000
ITEM 3	9.13	6.08	112	7.47	6.22	100	+1.97	+.135
ITEM 4	1.81	3.33	112	1.89	3.17	100	-0.17	-.010
ITEM 5	6.51	4.67	112	6.45	6.43	100	+0.08	+.000
ITEM 6	0.48	2.82	112	0.53	2.43	100	-0.13	-.010
ITEM 7	1.28	1.98	112	1.49	2.52	100	-0.69	-.048
ITEM 8	4.93	4.36	112	3.66	4.31	100	+2.13	+.145
ITEM 9	0.15	0.56	112	0.17	0.53	100	-0.24	-.017
ITEM 10	0.04	0.30	112	0.03	0.22	100	+0.16	+.010
ITEM 11	0.24	0.88	112	0.14	0.86	100	+0.84	+.058
ITEM 12	0.87	2.64	112	0.69	2.35	100	+0.54	+.037
ITEM 13	3.94	5.19	112	3.55	4.97	100	+0.55	+.039
ITEM 14	0.18	0.70	112	0.21	0.77	100	-0.31	-.022
ONE STUD	6.42	6.22	112	7.41	8.71	100	-0.96	-.066
SM GROUP	2.38	5.18	112	2.22	4.84	100	+0.24	+.017
LG GROUP	11.08	13.38	112	5.12	10.34	100	+3.60	+.241
WHOL CLAS	16.12	12.80	112	15.49	14.04	100	+0.34	+.022

appears to be toward instruction in larger groups of students. This pattern is advanced tentatively, for it is in direct contradiction to the observations of the MAC monitor onsite. (ETS, 1972a, p. 43)

Tables III-43 and III-44 present the Time 1 and Time 2 data for Oakland. Among the 14 student activity items, there were six relatively slight changes, all of which were in the direction opposite to what had been hypothesized. In Item 3, "Writing"; Item 5, "Listening"; Item 7, "Asking Questions"; Item 8, "Answering Questions"; and Item 11, "Initiating", the EXP school declined relative to the CON school, with the Time 1 EXP superiority either being reduced somewhat or with the CON school having a greater observed frequency at Time 2. Although none of the t- values is significant, it appears from an examination of the means for both schools at Times 1 and 2 that the major portion of the relative changes in these activities can be attributed to changes in the CON school. That is, the EXP school student activities are quite stable over time.

The stability of student activities in the EXP school is reflected in the results of the context of instruction. Relative to the CON school, the pattern of the EXP school's context of instruction does not change over time. Its initial superiority in the direction of individualized instruction, reflected by a range of t-values from -3.34 to +4.01 as one goes from "Whole Class" to "One Student", is nearly replicated in the range and magnitudes of the Time 2 t-values. However, the appearance of EXP school stability is a bit misleading. From an examination of the means for both schools, it is interesting to note that the EXP school's relative position was maintained while both schools moved toward individualized instruction.

The Student Observation data for San Antonio are presented in Tables III-45 and III-46. There were five changes among the 14 student activity items. In Item 2, "Referring to Reading" and Item 3, "Writing", the changes were relatively large and not in the direction hypothesized. For Item 2 an examination of the Times 1 and 2 means reveals nothing; but for Item 3 it can be seen that although the frequency of "Writing" remained stable in the EXP school, it nearly doubled in the CON school. In

(Text resumes on page III-116)

TABLE III-43

STUDENT CLASSROOM OBSERVATION

E/C, OAKLAND, TIME 1

ITEMS & CONTEXT	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\bar{X}_T$	S.D.	N	$\bar{X}$	S.D.	N	t-val.	r_{pb}
ITEM 1	4.04	3.97	72	4.28	6.29	95	-0.28	-.022
ITEM 2	0.17	1.01	72	0.02	0.14	95	+1.39	+.108
ITEM 3	8.08	7.82	72	7.41	7.70	95	+0.56	+.044
ITEM 4	1.10	2.58	72	0.86	2.76	95	+0.56	+.044
ITEM 5	7.72	6.29	72	9.13	7.54	95	-1.28	-.099
ITEM 6	0.08	0.71	72	0.60	3.17	95	-1.36	-.105
ITEM 7	0.89	1.49	72	0.83	1.31	95	+0.26	+.020
ITEM 8	3.12	3.68	72	1.81	2.42	95	+2.78	+.211
ITEM 9	0.29	0.86	72	0.33	0.96	95	-0.24	-.020
ITEM 10	0.14	1.18	72	0.02	0.14	95	+0.97	+.075
ITEM 11	0.49	1.22	72	0.21	0.96	95	+1.64	+.126
ITEM 12	1.35	3.48	72	0.96	3.20	95	+0.75	+.058
ITEM 13	5.37	3.80	72	5.63	5.03	95	-0.36	-.028
ITEM 14	2.40	3.64	72	2.57	3.79	95	-0.28	-.022
ONE STUD	9.97	7.09	72	6.23	4.95	95	+4.01	+.298
SM GROUP	8.74	9.08	72	6.64	7.74	95	+1.61	+.124
LG GROUP	8.12	8.68	72	6.89	8.59	95	+0.91	+.071
WHOL CLAS	8.69	10.18	72	14.78	12.67	95	-3.34	-.252

TABLE III-44

STUDENT CLASSROOM OBSERVATION
E/C, OAKLAND, TIME 2

ITEMS & CONTEXT	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
ITEM 1	5.24	6.28	68	4.58	6.00	100	+0.68	+.053
ITEM 2	0.09	0.38	68	0.02	0.20	100	+1.53	+.117
ITEM 3	8.26	8.29	68	10.22	9.84	100	-1.35	-.104
ITEM 4	1.24	2.89	68	1.07	3.04	100	+0.35	+.028
ITEM 5	9.56	8.68	68	7.88	9.37	100	+1.17	+.091
ITEM 6	1.18	4.46	68	1.33	4.94	100	-0.21	-.017
ITEM 7	1.03	1.74	68	1.27	1.97	100	-0.81	-.063
ITEM 8	2.65	3.71	68	2.31	3.43	100	+0.60	+.047
ITEM 9	0.24	0.65	68	0.34	0.84	100	-0.86	-.067
ITEM 10	0.22	0.69	68	0.16	0.63	100	+0.59	+.046
ITEM 11	0.71	2.49	68	0.76	3.69	100	-0.11	-.010
ITEM 12	0.29	1.84	68	0.22	1.74	100	+0.26	+.020
ITEM 13	5.35	5.32	68	4.53	4.22	100	+1.12	+.086
ITEM 14	0.72	2.11	68	0.75	1.71	100	-0.10	-.010
ONE STUD	14.10	9.88	68	7.96	7.45	100	+4.59	+.346
SM GROUP	4.01	5.39	68	3.18	4.49	100	+1.09	+.084
LG GROUP	5.26	8.00	68	4.08	8.37	100	+0.92	+.071
WHOL CLAS	13.43	10.77	68	20.07	11.95	100	-3.68	-.275

TABLE III-45

STUDENT CLASSROOM OBSERVATION
E/C, SAN ANTONIO, TIME 1

ITEMS & CONTEXT	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
ITEM 1	4.90	5.17	60	6.31	6.07	52	-1.33	-.125
ITEM 2	0.33	1.36	60	0.13	0.71	52	+0.95	+.090
ITEM 3	7.92	8.59	60	4.52	6.49	52	+2.33	+.217
ITEM 4	3.03	6.47	60	3.08	5.39	52	-0.04	-.000
ITEM 5	6.67	6.12	60	7.87	4.62	52	-1.16	-.110
ITEM 6	0.63	4.06	60	0.00	0.00	52	UNDEF	UNDEF
ITEM 7	1.17	1.94	60	1.29	2.09	52	-0.32	-.030
ITEM 8	3.93	5.67	60	5.75	5.89	52	-1.66	-.157
ITEM 9	0.65	1.90	60	0.29	0.89	52	+1.26	+.119
ITEM 10	0.08	0.33	60	0.02	0.14	52	+1.29	+.122
ITEM 11	0.13	0.79	60	0.21	1.27	52	-0.40	-.037
ITEM 12	2.43	5.05	60	3.00	7.11	52	-0.49	-.047
ITEM 13	3.83	4.31	60	3.27	3.53	52	+0.75	+.071
ITEM 14	0.25	0.91	60	0.23	1.00	52	+0.11	+.010
ONE STUD	7.18	7.01	60	8.65	7.16	52	-1.10	-.104
SM GROUP	5.40	6.54	60	4.04	5.71	52	+1.17	+.110
LG GROUP	8.97	9.49	60	3.62	7.59	52	+3.26	+.297
WHOL CLAS	14.12	12.95	60	19.15	12.00	52	-2.12	-.198

TABLE III-46

STUDENT CLASSROOM OBSERVATION
E/C, SAN ANTONIO, TIME 2

ITEMS & CONTEXT	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
ITEM 1	8.17	9.83	72	6.17	6.91	60	+1.33	+.116
ITEM 2	0.01	0.12	72	0.17	1.29	60	-1.00	-.087
ITEM 3	8.19	8.71	72	8.53	8.66	60	-0.22	-.020
ITEM 4	3.08	5.22	72	1.03	3.75	60	+2.54	+.218
ITEM 5	9.12	6.51	72	9.37	6.69	60	-0.21	-.017
ITEM 6	0.00	0.00	72	0.00	0.00	60	UNDEF	UNDEF
ITEM 7	1.46	1.78	72	1.80	2.18	60	-0.99	-.087
ITEM 8	2.53	3.99	72	3.97	5.04	60	-1.83	-.159
ITEM 9	0.32	1.02	72	0.18	0.54	60	+0.93	+.081
ITEM 10	0.03	0.17	72	0.00	0.00	60	UNDEF	UNDEF
ITEM 11	0.14	0.61	72	0.00	0.00	60	UNDEF	UNDEF
ITEM 12	1.25	3.93	72	0.47	1.37	60	+1.47	+.128
ITEM 13	0.71	1.69	72	1.42	2.39	60	-1.99	-.172
ITEM 14	0.22	0.17	72	0.25	0.82	60	-0.16	-.014
ONE STUD	10.00	9.93	72	10.80	12.58	60	-0.41	-.036
SM GROUP	3.01	5.42	72	3.12	6.39	60	-0.10	-.010
LG GROUP	3.81	6.42	72	2.43	5.80	60	+1.28	+.111
WHOL CLAS	18.89	14.14	72	17.33	14.99	60	+0.61	+.054

Item 1, "Reading"; Item 12, "Academic Game Playing"; and Item 13, "Nonproductive", the status of the EXP school improved relative to the CON school. Again, because the t-values at Time 2 are not large, there is no instance of a clear EXP school superiority; however, the trend of the t-value changes from Time 1 to Time 2 is slightly favorable to the EXP school. An examination of the means reveals little interest except for Item 1, "Reading"; there it can be seen that the EXP school nearly doubled its efforts from Time 1 to Time 2, whereas the CON school frequency was quite similar over the two observations.

The one change in the context of instruction in San Antonio was away from the direction hypothesized. The status of the EXP school relative to the CON school declined in "Whole Class", the t-values changing from -2.12 at Time 1 to +0.61 at Time 2. That is, while both schools increased the frequency of this context of instruction, the EXP school increase was greater, with a Time 1 mean of 14.11 and a Time 2 mean of 18.89.

Model-Based Comparisons--Student Classroom Observation Results

The first of the model-based comparisons to be examined is that for the Teacher-Only model. This comparison involves pooling the data from Cincinnati and Jacksonville. The results for these comparisons are presented in Tables III-47 and III-48. The results for the TO sites taken together resemble those for Jacksonville fairly closely. There are relative increases among the EXP students in the following activities: "Reading", "Writing", "Answering Questions", and "Playing Games". There are decreases among the EXP students (relative to the CON students) in the activities of "Refer to Reading", "Criticizing", and "Other Productive". The focus indexes show a shift away from a focus on "One Student" and a shift toward a focus on the "Large Group". Thus, these results provide seven confirmations of the predictions made: "Reading", "Writing", "Answering Questions" and "Referring to Reading" all increase; and "Playing Games", "Criticizing", and "Other Productive" decrease. There is one reversal of prediction, where there is a shift of the instructional focus

TABLE III-47

STUDENT CLASSROOM OBSERVATION
E/C, TEACHER ONLY, TIME 1

ITEMS & CONTEXT	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\bar{X}$	S.D.	N	$\bar{X}$	S.D.	N	t-val.	r_{pb}
ITEM 1	3.90	4.96	166	5.22	5.84	143	-2.16	-.122
ITEM 2	0.23	0.89	166	0.03	0.42	143	+2.38	+.135
ITEM 3	8.26	7.17	166	8.35	6.55	143	-0.12	-.000
ITEM 4	3.01	4.31	166	3.07	4.00	143	-0.12	-.000
ITEM 5	7.67	6.89	166	7.43	6.15	143	+0.33	+.020
ITEM 6	0.78	3.59	166	0.92	4.29	143	-0.28	-.017
ITEM 7	1.23	1.97	166	1.17	2.58	143	+0.23	+.014
ITEM 8	3.38	3.92	166	3.65	3.85	143	-0.61	-.035
ITEM 9	0.43	1.50	166	0.21	0.73	143	+1.62	+.092
ITEM 10	0.05	0.35	166	0.00	0.00	143	UNDEF	UNDEF
ITEM 11	0.28	0.75	166	0.19	0.80	143	+1.00	+.057
ITEM 12	1.03	2.93	166	1.45	4.94	143	-0.93	-.053
ITEM 13	4.27	4.62	166	3.34	4.84	143	+1.71	+.097
ITEM 14	1.23	3.46	166	0.22	1.03	143	+3.37	+.189
ONE STUD	9.10	9.27	166	7.05	7.92	143	+2.07	+.117
SM GROUP	4.31	6.53	166	3.62	6.29	143	+0.94	+.054
LG GROUP	5.74	8.99	166	10.11	12.43	143	-3.57	-.200
WHOL CLAS	16.87	13.45	166	14.57	14.33	143	+1.46	+.083

TABLE III-48

STUDENT CLASSROOM OBSERVATION
E/C, TEACHER ONLY, TIME 2

ITEMS & CONTEXT	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
ITEM 1	5.74	6.01	184	3.88	5.30	180	+3.14	+.163
ITEM 2	0.09	1.11	184	0.03	0.25	180	+0.70	+.037
ITEM 3	8.49	6.80	184	6.79	6.43	180	+2.44	+.127
ITEM 4	1.54	3.11	184	1.51	2.85	180	+0.10	+.000
ITEM 5	5.52	4.97	184	5.84	6.36	180	-0.54	-.028
ITEM 6	0.70	3.69	184	0.56	2.66	180	+0.41	+.022
ITEM 7	1.12	1.93	184	1.13	2.15	180	-0.04	-.000
ITEM 8	4.14	4.50	184	3.52	4.56	180	+1.29	+.068
ITEM 9	0.21	0.78	184	0.17	0.58	180	+0.62	+.033
ITEM 10	0.05	0.39	184	0.02	0.17	180	+1.20	+.063
ITEM 11	0.15	0.70	184	0.09	0.68	180	+0.72	+.037
ITEM 12	0.61	2.14	184	0.49	2.17	180	+0.53	+.028
ITEM 13	3.62	4.95	184	2.52	4.23	180	+2.29	+.120
ITEM 14	0.36	1.15	184	0.29	1.07	180	+0.60	+.032
ONE STUD	6.86	7.14	184	6.16	8.01	180	+0.88	+.047
SM GROUP	2.65	5.59	184	2.12	4.71	180	+0.98	+.051
LG GROUP	8.98	12.39	184	7.16	11.43	180	+1.46	+.076
WHOL CLAS	13.73	13.59	184	11.36	13.58	180	+1.67	+.087

toward large group instruction. In general, no conclusions should be drawn from these data for the TO model because it would appear that the Jacksonville results washed out the trends in Cincinnati.

The next of the model-based comparisons is that for the Parent-Teacher sites, Oakland and San Antonio. The results from this set of comparisons are presented in Tables III-49 and III-50. These data indicate that the San Antonio results dominate the Oakland results. Thus, the activities of "Reading", "Listening", and "Criticizing" decreases, for confirmations of hypotheses. However, "Playing Games" increases; the activities of "Writing", "Refer to Reading", "Asking Questions", "Answering Questions", and "Initiating" all decrease, which is contrary to what was hypothesized. There appears to be no significant trend to the context data. If anything, it would appear that the EXP schools tend not to use large group instruction quite as frequently as the CON schools, but even here the former's Time 1 advantage deteriorates (t-value from -3.53 to -1.86).

The third model-based comparison is that for the EXP students only in the PT sites versus the EXP students only in the TO sites. As mentioned previously, this comparison is subject to distortion from differences between sites in the kinds of behavior that is typical. However, it does give some hints as to possible differences in behavior promoted by the two models. The results are presented in Tables III-51 and III-52. In the Parent-Teacher Experimental schools, there is a relative increase in the following activities: "Reciting", "Listening", "Asking Questions", "Playing Games", and "Nonproductive Behavior". There are decreases in: "Refer to Reading", "Asking Questions", "Criticizing", and "Other Productive Behavior". This represents a confirmation of five predictions. However, there are also four reversals of prediction: "Nonproductive" increases; and "Referring to Reading", "Asking Questions" and "Other Productive Behavior" decrease. The context indexes for the PT-E versus TO-E comparisons indicate a relative increase in focus on "One Student" and on "Whole Class", and a relative decrease in focus on "Small Group" and "Large Group". This pattern could be interpreted in a variety of ways.

(Text resumes on page III-124)

TABLE III-49

STUDENT CLASSROOM OBSERVATION
E/C, PARENT TEACHER ONLY, TIME 1

ITEMS & CONTEXT	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
ITEM 1	4.43	4.56	132	5.00	6.26	147	-0.86	-.052
ITEM 2	0.24	1.18	132	0.06	0.44	147	+1.73	+.103
ITEM 3	8.01	8.15	132	6.39	7.40	147	+1.74	+.104
ITEM 4	1.98	4.84	132	1.65	4.02	147	+0.62	+.037
ITEM 5	7.24	6.21	132	8.68	6.67	147	-1.86	-.111
ITEM 6	0.33	2.79	132	0.39	2.56	147	-0.17	-.010
ITEM 7	1.02	1.71	132	0.99	1.64	147	+0.11	+.000
ITEM 8	3.49	4.69	132	3.20	4.41	147	+0.53	+.032
ITEM 9	0.45	1.43	132	0.31	0.93	147	+0.99	+.059
ITEM 10	0.11	0.90	132	0.02	0.14	147	+1.24	+.075
ITEM 11	0.33	1.06	132	0.21	1.07	147	+0.90	+.054
ITEM 12	1.84	4.28	132	1.68	5.02	147	+0.29	+.017
ITEM 13	4.67	4.10	132	4.80	4.68	147	-0.23	-.014
ITEM 14	1.42	2.96	132	1.74	3.29	147	-0.84	-.051
ONE STUD	8.70	7.17	132	7.09	5.92	147	+2.06	+.123
SM GROUP	7.22	8.17	132	5.72	7.18	147	+1.63	+.097
LG GROUP	8.51	9.03	132	5.73	8.37	147	+2.66	+.158
WHOL CLAS	11.16	11.79	132	16.33	12.57	147	-3.53	-.208

TABLE III-50

STUDENT CLASSROOM OBSERVATION
E/C, PARENT TEACHER ONLY, TIME 2

ITEMS & CONTEXT	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
ITEM 1	6.74	8.39	140	5.17	6.38	160	+1.83	+.106
ITEM 2	0.05	0.28	140	0.07	0.81	160	-0.35	-.020
ITEM 3	8.23	8.48	140	9.59	0.42	160	-1.31	-.075
ITEM 4	2.19	4.33	140	1.06	3.31	160	+2.55	+.146
ITEM 5	9.34	7.62	140	8.44	8.47	160	+0.96	+.056
ITEM 6	0.57	3.15	140	0.83	3.95	160	-0.62	-.036
ITEM 7	1.25	1.76	140	1.47	2.06	160	-0.98	-.057
ITEM 8	2.59	3.84	140	2.93	4.17	160	-0.74	-.042
ITEM 9	0.28	0.86	140	0.28	0.75	160	-0.03	-.000
ITEM 10	0.12	0.50	140	0.10	0.50	160	+0.37	+.022
ITEM 11	0.41	1.80	140	0.47	2.93	160	-0.21	-.014
ITEM 12	0.79	3.12	140	0.31	1.61	160	+1.68	+.097
ITEM 13	2.96	4.53	140	3.36	3.93	160	-0.81	-.047
ITEM 14	0.46	1.70	140	0.56	1.46	160	-0.54	-.032
ONE STUD	11.99	10.08	140	9.02	9.75	160	+2.59	+.148
SM GROUP	3.50	5.41	140	3.16	5.27	160	+0.56	+.032
LG GROUP	4.51	7.24	140	3.46	7.53	160	+1.23	+.071
WHOL CLAS	16.24	12.87	140	19.04	13.20	160	-1.86	-.107

TABLE III-51

STUDENT CLASSROOM OBSERVATION
PARENT TEACHER-E/TEACHER ONLY-E, TIME 1

ITEMS & CONTEXT	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
ITEM 1	4.43	4.56	132	3.90	4.96	166	+0.96	+.056
ITEM 2	0.24	1.18	132	0.23	0.89	166	+0.11	+.000
ITEM 3	8.01	8.15	132	8.26	7.17	166	-0.28	-.017
ITEM 4	1.98	4.84	132	3.01	4.31	166	-1.95	-.113
ITEM 5	7.24	6.21	132	7.67	6.89	166	-0.56	-.033
ITEM 6	0.33	2.79	132	0.78	3.59	166	-1.18	-.069
ITEM 7	1.02	1.71	132	1.23	1.97	166	-1.01	-.059
ITEM 8	3.49	4.69	132	3.38	3.92	166	+0.23	+.014
ITEM 9	0.45	1.43	132	0.43	1.50	166	+0.12	+.000
ITEM 10	0.11	0.90	132	0.05	0.35	166	+0.78	+.046
ITEM 11	0.33	1.06	132	0.28	0.75	166	+0.46	+.026
ITEM 12	1.84	4.28	132	1.03	2.93	166	+1.94	+.112
ITEM 13	4.67	4.10	132	4.27	4.62	166	+0.80	+.047
ITEM 14	1.42	4.96	132	1.23	3.46	166	+0.50	+.028
ONE STUD	8.70	7.17	132	9.10	9.27	166	-0.40	-.022
SM GROUP	7.22	8.17	132	4.31	6.53	166	+3.42	+.195
LG GROUP	8.51	9.03	132	5.74	8.99	166	+2.63	+.151
WHOL CLAS	11.16	11.79	132	16.87	13.45	166	-3.85	-.218

TABLE III-52

STUDENT CLASSROOM OBSERVATION
PARENT TEACHER-E/TEACHER ONLY-E, TIME 2

ITEMS & CONTEXT	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\bar{X}$	S.D.	N	$\bar{X}$	S.D.	N	t-val.	r_{pb}
ITEM 1	6.74	8.39	140	5.74	6.01	184	+1.25	+.069
ITEM 2	0.05	0.28	140	0.09	1.11	184	-0.39	-.022
ITEM 3	8.23	8.48	140	8.49	6.80	184	-0.31	-.017
ITEM 4	2.19	4.-3	140	1.54	3.11	184	+1.57	+.087
ITEM 5	9.34	7.62	140	5.52	4.97	184	+5.45	+.290
ITEM 6	0.57	3.15	140	0.70	3.69	184	-0.33	-.017
ITEM 7	1.25	1.76	140	1.12	1.93	184	+0.62	+.035
ITEM 8	2.59	3.84	140	4.14	4.50	184	-3.27	-.179
ITEM 9	0.28	0.86	140	0.21	0.78	184	+0.73	+.040
ITEM 10	0.12	0.50	140	0.05	0.39	184	+1.36	+.075
ITEM 11	0.41	1.80	140	0.15	0.70	184	+1.84	+.102
ITEM 12	0.79	3.12	140	0.61	2.14	184	+0.59	+.033
ITEM 13	2.96	4.53	140	3.62	4.95	184	-1.23	-.069
ITEM 14	0.46	1.70	140	0.36	1.15	184	+0.63	+.035
ONE STUD	11.99	10.08	140	6.86	7.14	184	+5.36	+.286
SM GROUP	3.50	5.41	140	2.65	5.59	184	+1.38	+.077
LG GROUP	4.51	7.24	140	8.98	12.39	184	-3.80	-.207
WHOL CLAS	16.24	12.87	140	13.73	13.59	184	+1.68	+.093

The last of the model-based comparisons for the student observation data is the composite of all students in the EXP schools versus all students in the CON schools. These results are presented in Tables III-53 and III-54. The comparison shows a relative increase for the EXP students in three activities. These are: "Readin ", "Listening", and "Academic Gaming". There are relative decreases in three other activities; "Refer to Reading", "Asking Questions", "Criticizing", and "Other Productive Behavior". This represents a confirmation of three predictions: "Reading" increases; "Academic Gaming" increases; and "Criticizing" decreases. There are three changes opposite to the direction predicted: "Referring to Reading", "Asking Questions", and "Other Productive Behavior" decrease. The focus indexes for this overall comparison show no change from the first to second time point, with the EXP students receiving more individualized instruction relative to the CON students.

TABLE III-53

STUDENT CLASSROOM OBSERVATION
E/C, ALL, TIME 1

ITEMS & CONTEXT	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
ITEM 1	4.13	4.78	298	5.11	6.05	290	-2.17	-.089
ITEM 2	0.23	1.03	298	0.05	[illegible].43	290	+2.86	+.117
ITEM 3	8.15	7.60	298	7.36	7.05	290	+1.31	+.054
ITEM 4	2.55	4.57	298	2.35	4.07	290	+0.57	+.024
ITEM 5	7.48	6.59	298	8.06	6.44	290	-1.08	-.045
ITEM 6	0.58	3.26	298	0.64	3.52	290	-0.22	-.010
ITEM 7	1.14	1.86	298	1.08	2.15	290	+0.33	+.014
ITEM 8	3.43	4.27	298	3.42	4.14	290	+0.02	+.000
ITEM 9	0.44	1.47	298	0.26	0.84	290	+1.83	+.075
ITEM 10	0.08	0.65	298	0.01	0.10	290	+1.81	+.075
ITEM 11	0.30	0.90	298	0.20	0.95	290	+1.29	+.053
ITEM 12	1.39	3.61	298	1.57	4.97	290	-0.50	-.020
ITEM 13	4.45	4.39	298	4.08	4.81	290	+0.97	+.040
ITEM 14	1.32	3.24	298	0.99	2.57	290	+1.35	+.056
ONE STUD	8.92	8.39	298	7.07	[illegible].97	290	+2.91	+.119
SM GROUP	5.60	7.43	298	4.68	6.82	290	+1.55	+.064
LG GROUP	6.97	9.10	298	7.89	10.77	290	-1.13	-.047
WHOL CLAS	14.34	13.03	298	15.46	13.47	290	-1.02	-.042

TABLE III-54

STUDENT CLASSROOM OBSERVATION
E/C, ALL, TIME 2

ITEMS & CONTEXT	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
ITEM 1	6.18	7.14	324	4.49	5.86	340	+3.34	+.128
ITEM 2	0.07	0.85	324	0.05	0.58	340	+0.37	+.014
ITEM 3	8.38	7.56	324	8.11	8.09	340	+0.44	+.017
ITEM 4	1.82	3.70	324	1.29	3.08	340	+1.99	+.077
ITEM 5	7.17	6.53	324	7.06	7.53	340	+0.19	+.010
ITEM 6	0.65	3.47	324	0.69	3.34	340	-0.16	-.000
ITEM 7	1.18	1.86	324	1.29	2.11	340	-0.73	-.028
ITEM 8	3.47	4.29	324	3.24	4.38	340	+0.66	+.026
ITEM 9	0.24	0.82	324	0.22	0.67	340	+0.35	+.014
ITEM 10	0.08	0.44	324	0.06	0.37	340	+0.87	+.033
ITEM 11	0.26	1.30	324	0.27	2.08	340	-0.08	-.000
ITEM 12	0.69	2.61	324	0.41	1.93	340	+1.58	+.061
ITEM 13	3.34	4.78	324	2.91	4.11	340	+1.23	+.048
ITEM 14	0.41	1.41	324	0.42	1.27	340	-0.13	-.000
ONE STUD	9.08	8.89	324	7.51	8.97	340	+2.27	+.069
SM GROUP	3.02	5.52	324	2.61	5.00	340	+1.00	+.039
LG GROUP	7.05	10.70	324	5.42	9.95	340	+2.04	+.079
WHOL CLAS	14.81	13.32	324	14.97	13.92	340	-0.15	-.000

Student Classroom Observation Results--
Discussion

A summary of the results of the analysis of the Student Classroom data that would contribute either to a better understanding of the impact of incentives or to any organized body of knowledge about teaching is impossible. Table III-55 summarizes the results of the hypotheses, by site. There were a total of 72 opportunities to test the hypotheses concerning the 14 student activities and four context variables. In only 28 out of these 72 cells was any change noticed. Of this 37% in which change did occur, only 13, or 46%, were in the direction hypothesized. Overall, theresults are at best unclear.

One generalization that does seem warranted, however, is that the activity of "Reading" increases differentially in the EXP schools as compared with the CON schools. The pattern is reflected by an increase in t-values greater than 1.00 in three of the sites, and an increase of 0.97 in Oakland. However, the source of these increases, in terms of the actual frequencies of the activity "Reading" in each of the schools at each of the times, is not so simple as might be thought. For instance, in Cincinnati, the "Reading" results are produced through the occurrence of a slight increase in the frequency of "Reading" at the EXP school, and a larger decrease in frequency of the same activity at the CON school. Likewise, in Jacksonville, the EXP school shows a considerable increase in frequency of "Reading", and the CON school shows a sizable decrease. Thus, in both of these cases, the apparent magnitude of the effect is determined partly by a change in the behavior at the CON school over the interval between December and May, as well as by the changes in the behavior at the EXP school.

For Oakland, the pattern is somewhat different. There is an increase in the frequency of the "Reading" activity at the EXP school, and no change worth noting at the CON school. A similar pattern prevails at San Antonio, where the frequency of "Reading" as an activity increases substantially at the EXP school, and shows little change at the CON school.

TABLE III-55

STUDENT CLASSROOM OBSERVATION RESULTS
COMPARISON OF HYPOTHESES AND RESULTS

ITEM/ CONTEXT	ACTIVITY	RESULTS[1]				
		HYP.	CIN	JAX	OAK	SAN
1	Reading	E > C	+	+	+	+
2	Referring to Reading	E > C				-
3	Writing	E > C		+	-	-
4	Reciting	E > C				
5	Listening	E < C			-	
6	Being Tested	E > C	+			
7	Asking Questions	E > C	+	-	-	
8	Answering Questions	E > C		+	-	
9	Criticizing	E < C		+		
10	Complimenting	E > C				
11	Initiating	E > C			-	
12	Academic Games	E < C		+		+
13	Nonproductive	E < C	-		-	+
14	Productive	E > C	-	-		
	One Student		+	-		
	Small Group					
	Large Group			-		
	Whole Class					-

Key:

+ = Results are in the direction hypothesized.
- = Results are not in the direction hypothesized.

Thus, the source of the apparently uniform results for "Reading" across all four sites is found to be different in the two sites which experienced the TO model from that in the two sites which experienced the PT model. For this particular pattern, a number of substantive explanations are possible. For instance, it might be argued that in the two Parent-Teacher sites, some form of parent pressure was being exerted which led to the increased use of class time for reading. The difficulty with this possibility is that it cannot be directly evaluated with the data that has been collected in the project. The Parent data does not bear directly on this question, and in any case is not complete. Moreover, one further problem involved in attributing the observed differences to parent pressure is that in the two Parent-Teacher sites, most of the parents were not aware of the existence of a project, much less its exact provisions, until quite late in the year. Of course, this could still have affected the Time 2 observations, which were made in May. But, this explanation must remain as a plausible speculation.

More generally, there are a number of problems in drawing inferences from the classroom observations of student behavior. For example, an alternate explanation for the increase of class time spent in reading is simply that this was a direct effect of the offering of incentives to the teachers for the reading gains made by their students. That is, the increase in reading could be explained simply as evidence that the offer of incentives had brought about a change in the proportion of classroom time devoted to different kinds of learning. Since reading now was a kind of learning for which incentives were being offered, the amount of time spent on reading was increased. This explanation, like the one mentioned previously, is plausible but not testable with the data at hand. That is, the increase does not occur uniformly across the four sites, and is in fact partly due to a decrease in the frequency of reading at the CON schools.

With the exception of "Reading", there are no consistent patterns or trends in the data that bear directly on the impact of incentives. "Writing" decreases in Oakland and San Antonio, but it increases in Jacksonville.

"Individualized Instruction" increases in Cincinnati, but decreases in Jacksonville. "Nonproductive" behavior increases in Cincinnati and Oakland, but only in Cincinnati is there a concomitant decline in "Other Productive" behavior. Meanwhile, in San Antonio, "Nonproductive" behavior decreases, as predicted. Across the four sites, then, at least as indicated by this data, influences other than the offer of incentives appear to be working.

Within the sites, there are anomalies that appear not to be in accord with the conventional wisdom about teaching and learning. In Cincinnati, for example, "Reading", "Being Tested", and "Asking Questions" all confirm the hypotheses. One might take this to mean that there was a tightening up in the EXP school, a vigorous focus upon the subject matter and the objectives of the project. This assumption is made at great risk, however, because there was a simultaneous increase in "Nonproductive" behavior and decrease in "Productive" behavior. In Jacksonville, "Asking Questions", generally attributed to openness in the classroom, declined while other indices of openness increased. Furthermore, if "Asking Questions" is an isolated deviation from a trend toward more interaction and spontaneity in the classroom, how does one explain the trend away from "One Student" toward "Large Group" instruction?

To further complicate matters, there is some congruence between the individual site data discussed here and the MAC observations and reports (Education Turnkey Systems, Inc., 1972a & b). Some hint of the relative influence of the absence of the EXP school's principal reported by MAC can be gleaned from this data. In Oakland, it appears that the erosion of the EXP school principal's authority along with maintenance of the CON school principal's strength and direction are reflected in the relative deterioration of the EXP school. However, although MAC made much of the principal's authority, less was made of the fact that a significant percentage of the EXP school faculty simply refused to acknowledge the existence of the project. They claimed, from time to time quite vehemently, that they were in the project under protest and would therefore do nothing different. The data confirm their promise.

In addition to the variety of possible substantive explanations, there are also several possible sources of methodological complications involved in this observation data. A brief list of these will indicate the kinds of issues that are involved:

- The categories are ipsative, and so a change in one category frequency will tend to imply changes in the frequencies in other categories as well.
- The instrument is of unknown general reliability.
- Differences between observers in terms of the way particular events are coded could contribute to differences between sites, although not to differences within sites.
- Differences could emerge in the skill of the observers between the first and the second wave of observations, and these differences would likely lead to changes in the pattern of frequencies that were coded.
- The use of only the school averages for analysis, while desirable as a means of dampening out random fluctuations, also may conceal considerable differences between individual teachers as to the kinds of interaction patterns observed.
- The interaction patterns observed may in fact be highly variable from situation to situation, and the totals found may well be largely the result of random sampling of the situational differences.

These potential sources of method artifacts were recognized, but could not be entirely neutralized, given the circumstances of the project. Coupled with the complex pattern of results observed at each site, they make it necessary to avoid drawing inferences from this data as to the specific kinds of substantive changes that might be expected from the implementation of an incentives model. The frequency of reading time in the EXP schools relative to the CON schools has been noted, but the reader should keep in mind the currents and eddies in the data that contributed to this outcome.

Analysis of Teacher Attitudes and Behavior

Introduction

Information concerning teacher reactions to the incentives models, as reflected in their expressed attitudes and selected behaviors, was collected by TAC via three instruments: (1) a Teacher Questionnaire, (2) a Teacher Interview, and (3) a Structured Classroom Observation Form. Each of these instruments was administered at two times during the year.

The form for the Teacher Questionnaire was not given clearance by officials at OMB and OE until February, 1972, and so was first administered at that time. The number of teachers who completed the Teacher Questionnaire at the time it was first administered, denoted Time 1, was 103. The number of teachers responding to the Time 2 Teacher Questionnaire, administered in May, was 81. There were 151 teachers defined as eligible to fill out these Questionnaires, and so the response rate for the Time 1 Questionnaire was 68.2%, and the response rate for the Time 2 Questionnaire was 53.6%.

The Teacher Interview data collection was carried out at two times during the year. The first wave of interviews with the teachers was conducted during November and December, and the second wave was conducted during May. At each of these two times, interviews were successfully completed with all 151 eligible teachers.

The Structured Classroom Observation of teachers was carried out concurrently with the similar observation of student behavior described in an earlier section. These Structured Observations took place at two times during the year, namely December and May. They made use of repeated sampling of class sessions at each of the two times so as to avoid the possible biases that might arise from carrying out observation during one session which might be atypical. The observations were recorded by means of 14 categories. The coding scheme and the results of these observations are discussed in the section following this one.

The data from the Teacher Questionnaires and from the Teacher Interviews are similar in that all of it concerns a teacher's verbally expressed attitudes on topics which might well be influenced by experience with the incentives model. Accordingly, the results from both of these instruments will be reviewed and discussed together in this section.

Before proceeding, it should be pointed out that, as with the Student Questionnaire and Student Interview data, the fact that the first administration of the Questionnaire occurred so late in the year creates a serious problem in comparing the results of the Questionnaire data with the results of the Interview data. In addition, the relatively incomplete number of responses obtained to the Teacher Questionnaire (which was distributed via a mail-back method), and the decrease in response rate between the two waves of the Questionnaire, introduce additional possible sources of bias into the results.

Index Construction

The data obtained from the 28 items of the Teacher Questionnaire were summarized by means of 11 indexes. These indexes were constructed in a manner similar to that used with the student data. That is, the questions had been developed with particular hypothesized processes in mind, and so the items could be combined into indexes. Those questions which dealt with a single topic were combined as components for an index. The responses made by each teacher to those items were added together to provide the index score for that teacher on that index. The concepts around which the Teacher Indexes were built included: attitudes toward paying monetary incentives to teachers and parents; attitudes toward the school; attitudes toward the students; perceived quality of interpersonal relations among the teachers; and preferences and opinions about various teaching methods. The name of each index and the list of items which compose it is given in Table III-56.

Similarly, the data obtained from the 19 items of the Teacher Interview were also summarized in a set of four indexes. The interview was relatively brief, and the smaller number of items precluded development of a large

TABLE III-56

INDEXES CONSTRUCTED FROM THE TEACHER QUESTIONNAIRE

1. Teacher Attitude toward Students

Question No.	
5	My students' motivation to learn is...(very poor) ...(very high)
6	My students' academic ability is...(very limited)...(very capable)
7	My students behave...(very poorly)...(very well)
26	My attitude toward my pupils is...(strongly negative)...(strongly positive)

2. Teacher Attitude toward School

Question No.	
4	My school...(could stand alot of improvement)...(could serve as a model to others)
12	My feeling about transferring to another school is...(very favorable)...(very opposed)
16	My attitude toward my school is...(strongly negative)...(strongly positive)

3. Teacher Attitude toward Incentives to Teachers

Question No.	
9	The offer of incentives to a teacher based on the achievement of his or her students is...(unlikely to increase achievement)...(likely to increase achievement)
21	The offer of incentives to a teacher based on the achievement of his or her students is...(not proper)...(proper)

4. Teacher Attitude toward Incentives to Parents

Question No.	
15	The offer of incentives to parents based on the achievement of their child's class is...(not proper)...(proper)
24	The offer of incentives to parents based on the achievement of their child's class is...(unlikely to increase achievement)...(likely to increase achievement)

5. Teacher Attitude toward Relaxed Discipline

Question No.	
14	Letting children move around in the classroom and talk to each other...(prevents learning)...(fosters learning)
25	Strict discipline in the classroom is an important part of a child's education...(strongly agree)...(strongly disagree)

-Continued

TABLE III-56 (Cont'd.)

6. Teacher Perception of Principal's Attitude toward Innovations

Question No.

10 If I want to use new teaching methods or materials, my principal is likely to...(oppose me)...(assist me)

19 My principal's attitude toward new teaching methods and materials is...(negative)...(positive)

7. Teacher Perception of Cooperation Among Faculty

Question No.

17 Faculty meetings at my school are...(counterproductive)...(productive)

18 Efforts of the faculty at my school to assist one another are ...(counterproductive)...(productive)

8. Adult Relations in the School

Question No.

This index is the sum of indexes 6 and 7.

9. Teacher Attitude toward Individualized Instruction

Question No.

22 If a child is allowed to proceed at his own rate during a year's time, he is likely to...(learn less)...(learn more)

28 Individualized and self-paced instructions are likely to be...(very ineffective)...(very effective)

10. Teacher Attitude toward Parents

Question No.

11 The role most of my pupils' parents play in their children's education is...(of no significance)...(of great significance)

20 My attitude toward my children's parents is...(strongly negative)...(strongly positive)

11. Teacher Attitude toward Peer Tutoring

Question No.

23 My feelings about having faster pupils tutor slow pupils in my classroom are...(opposed)...(favorable)

27 Having faster pupils tutor slower pupils is likely to be...(very ineffective)...(very effective)

number of indexes. The names of the indexes constructed from the Teacher Interview, and the items composing each, are given in Table III-57.

Single-Site Comparisons--Teacher Questionnaire and Interview Results

The first of the individual sites to be examined is Cincinnati. The results for Cincinnati are presented in Table III-58. The pattern of results is generally that there are few major changes, but that the changes which do occur are in a negative direction. For instance, of the 11 indexes constructed from the Teacher Questionnaire, only two show a substantial shift in the t-values from Time 1 to Time 2. These are: "Attitude toward Students" and "Attitude Toward Individualized Instruction". For both of these indexes, the shifts are in the direction opposite to what had been predicted by the hypothesis. The t-value at Time 1 for the "Attitude toward Students" index is +1.30, and the t-value on the same index at Time 2 is -0.32. This shift is one that may be due to regression effects, and so there is some ambiguity surrounding it. Similarly, on the index for "Attitude toward Individualized Instruction", the t-value changes from +2.51 to +1.46. This shift also could be attributed to a regression toward the mean. However, since these are the only changes in t-value larger than 1.00, it is clear that there is no positive impact of the incentives model at Cincinnati, as far as teacher attitudes expressed on the Questionnaire are concerned. When it is considered that the incomplete response rate probably led to a differential exclusion of teachers who are unfavorable to the project, this conclusion becomes even stronger.

The general trend of the results, though not the specific attitudes, are similar when the data from the Teacher Interview are examined. Again, only two of the indexes from the Teacher Interview show substantial change in the t-values, and both of these changes are in a direction opposite to that hypothesized. The two indexes which show change are: "Attitude toward Modern Techniques" and "Relations with Parents". Here, too, the changes observed are in a direction that might be predicted from the operation of a regression effect. For instance, the t-value for "Attitude toward Modern Techniques" at Time 1 is +2.35, and at Time 2 the t-value

(Text resumes on page III-141)

TABLE III-57

INDEXES CONSTRUCTED FROM THE TEACHER INTERVIEW

1. Teacher Attitude toward Incentives Idea and Project

Question No.	
1	Do you think the payment of incentives to teachers is a good thing
2	Do you think that the payment of incentives to parents is a good thing
18	What are your feelings toward the incentives project being tried in your school district

2. Teacher Attitude toward Modern Techniques

Question No.	
4	Have you made use of any techniques to individualize the curriculum you use in the classroom
5	How often in the last two months have you requested special materials (books, films, etc.) not normally supplied
8	In a normal school week do you use peer tutoring as a classroom teaching technique
12	Do you feel that the most important thing in the classroom should be discipline

3. Teacher Relations with Parents

Question No.	
10	What percent of your students' parents have visited you in the last two months
11	Do you feel you get on well with the parents of the children you teach
13	How do you feel about the ability of the parents of the children you teach to help their children with their homework

3. Teacher Dedication to Teaching

Question No.	
6	Are you satisfied with the profession of teaching for you as an individual
7	In the last two weeks have you ever considered transferring out of the school you are presently teaching in
9	Do you enjoy your students as children
14	How many hours during a normal week do you spend on lesson preparation out of school

TABLE III-58-1

TEACHER QUESTIONNAIRE
E/C, CINCINNATI, TIME 1

ATTITUDES	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\bar{X}$	S.D.	N	$\bar{X}$	S.D.	N	t-val.	r_{pb}
ATT TO STDNTS	12.70	2.41	10	11.00	2.76	6	+1.30	+.328
ATT TO SCHOOL	9.20	2.53	10	10.17	1.17	6	-0.87	-.227
INCEN TO TEACH	5.00	1.48	11	5.00	2.37	6	+0.00	+.000
INCEN TO PARNT	6.09	2.63	11	3.17	1.17	6	+2.56	+.552
RELAXED DISC	6.00	1.41	11	6.67	1.63	6	-0.83	-.222
ATT PRIN RE INNOV	9.00	1.10	11	8.83	0.98	6	+0.31	+.080
COOP FACULTY	7.82	1.54	11	8.50	1.22	6	-0.93	-.234
ADULT REL	16.82	1.89	11	17.33	1.75	6	-0.55	-.014
INDIV INST	8.91	1.04	11	7.67	0.82	6	+2.51	+.544
ATT RE PARNT	5.90	1.20	10	6.00	1.41	6	-0.15	-.040
PEER TUTORING	8.91	0.83	11	8.33	1.51	6	+1.03	+.257

TABLE III-58-2

TEACHER QUESTIONNAIRE
E/C, CINCINNATI, TIME 2

ATTITUDES	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
ATT TO STDNTS	13.50	3.39	6	14.33	4.16	3	-0.32	-.122
ATT TO SCHOOL	9.37	2.83	8	10.33	1.53	3	-0.55	-.179
INCEN TO TEACH	4.50	1.85	8	4.33	2.08	3	+0.13	+.044
INCEN TO PARNT	5.86	2.85	7	3.00	1.73	3	+1.58	+.488
RELAXED DISC	6.14	1.07	7	6.67	2.08	3	-0.54	-.189
ATT PRIN RE INNOV	7.87	1.96	8	8.33	0.58	3	-0.31	-.128
COOP FACULTY	7.50	1.20	8	8.67	0.58	3	-1.58	-.467
ADULT REL	15.37	2.56	8	17.00	0.00	3	UNDEF	UNDEF
INDIV INST	8.14	1.21	7	7.00	1.00	3	+1.42	+.449
ATT RE PARNT	6.37	1.60	8	7.33	1.53	3	-0.89	-.286
PEER TUTORING	8.86	1.21	7	7.33	1.15	3	+1.84	+.545

TABLE III-58-3

TEACHER INTERVIEW
E/C, CINCINNATI
TIME 1

ATTITUDES	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
INCENTIVES	9.07	3.27	14	7.44	2.50	16	+1.55	+.281
MODERN TECH	11.93	2.62	14	9.50	2.99	16	+2.35	+.406
PARENTS	13.00	1.57	14	11.50	2.00	16	+2.26	+.393
DEDICATION	18.57	1.02	14	17.69	2.68	16	+1.16	+.214

TIME 2

ATTITUDES	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
INCENTIVES	8.38	2.66	13	6.40	2.56	15	+2.01	+.367
MODERN TECH	12.38	2.75	13	11.69	3.79	16	+0.55	+.106
PARENTS	9.69	2.17	13	10.33	2.19	15	-0.77	-.150
DEDICATION	18.00	1.83	13	17.81	1.83	16	+0.27	+.053

is +0.55. Similarly, the t-value for the "Relations with Parents" index at Time 1 is +2.26, and at Time 2 is -0.77. It is worth noting, however, that the shift in "Relations with Parents" occurs primarily because of a large decrease in the average obtained by teachers in the EXP school from Time 1 to Time 2. The shift on the index of "Attitude toward Modern Techniques", on the other hand, is accomplished through an increase in the average among the teachers in the CON school.

In the second of the sites, Jacksonville, there are three Teacher Questionnaire indexes which show some change. These are: "Attitude toward Students", "Attitude toward Relaxed Discipline", and "Attitude toward Peer Tutoring". The results for Jacksonville are presented in Table III-59. All of these changes are in the hypothesized direction, and none of them is likely to be due to regression effects. That is, in all three cases, the size of the gap increases from Time 1 to Time 2, and the direction is in the favor of the EXP school. While these differences seem to be real, and seem also to reflect a general tendency to move toward a warmer and more relaxed classroom style, it is noteworthy that none of the indexes referring to attitudes toward the incentives project, or adult relations in the school, showed any sizable shifts.

In view of the pattern observed from the Teacher Questionnaire indexes in Jacksonville, one would think it highly probable that the index for "Attitude toward Modern Techniques" drawn from the Teacher Interview, also would show a shift in the predicted direction. However, this is not the case. The t-value for the Interview-based index of "Attitude toward Modern Techniques" at Time 1 is +1.42, and at Time 2 the corresponding t-value is +1.46. In fact, none of the Interview-based indexes shows any change of note for Jacksonville. This lack of confirmation is somewhat puzzling, but does not weaken the general conclusion based upon the Questionnaire results that there probably was an improvement in classroom warmth in Jacksonville at the EXP schools.

Inspection of the actual means for the three Questionnaire-based indexes which showed a shift in the predicted direction between Time 1 and Time 2 does lead to some qualification of the results. For the "Attitude

(Text resumes on page III-145)

TABLE III-59-1

TEACHER QUESTIONNAIRE
E/C, JACKSONVILLE, TIME 1

ATTITUDES	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\bar{X}$	S.D.	N	$\bar{X}$	S.D.	N	t-val.	r_{pb}
ATT TO STDNTS	13.35	3.22	17	11.92	4.44	13	+1.02	+.190
ATT TO SCHOOL	9.56	3.83	16	8.31	3.40	13	+0.92	+.175
INCEN TO TEACH	6.71	2.87	17	5.62	2.66	13	+1.06	+.197
INCEN TO PARNT	5.94	2.41	17	5.69	2.87	13	+0.26	+.049
RELAXED DISC	6.47	1.84	17	5.85	2.51	13	+0.79	+.147
ATT PRIN RE INNOV	6.76	3.40	17	8.08	1.66	13	-1.28	-.235
COOP FACULTY	6.94	2.62	16	6.00	2.16	13	+1.03	+.195
ADULT REL	14.00	5.72	16	14.08	3.59	13	-0.04	-.010
INDIV INST	8.88	1.69	17	8.38	1.50	13	+0.84	+.156
ATT RE PARNT	7.53	1.42	17	6.23	2.09	13	+2.03	+.358
PEER TUTORING	8.94	1.75	17	8.54	1.94	13	+0.60	+.112

TABLE III-59-2

TEACHER QUESTIONNAIRE
E/C, JACKSONVILLE, TIME 2

ATTITUDES	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
ATT TO STDNTS	14.23	3.35	13	9.00	4.64	5	+2.68	+.556
ATT TO SCHOOL	8.75	4.37	12	6.20	3.49	5	+1.15	+.285
INCEN TO TEACH	6.08	2.36	13	4.83	2.48	6	+1.05	+.247
INCEN TO PARNT	5.62	2.40	13	4.83	2.56	6	+0.65	+.155
RELAXED DISC	6.38	1.94	13	4.67	1.63	6	+1.88	+.414
ATT PRIN RE INNOV	6.15	3.48	13	8.00	1.67	6	-1.22	-.284
COOP FACULTY	6.08	3.01	13	5.17	2.23	6	+0.66	+.157
ADULT REL	12.23	6.38	13	13.17	3.71	6	-0.33	-.080
INDIV INST	9.08	1.93	13	8.00	1.67	6	+1.17	+.273
ATT RE PARNT	7.00	1.08	13	5.67	1.97	6	+1.93	+.424
PEER TUTORING	9.69	0.63	13	8.83	1.33	6	+1.95	+4.27

TABLE III-59-3

TEACHER INTERVIEW
E/C, JACKSONVILLE
TIME 1

ATTITUDES	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
INCENTIVES	10.62	3.06	26	9.57	3.49	21	+1.09	+.161
MODERN TECH	13.04	3.49	26	11.67	3.01	21	+1.42	+.208
PARENTS	10.58	2.10	26	10.48	2.23	21	+0.16	+.024
DEDICATION	17.27	2.39	26	17.00	3.19	21	+0.33	+.049

TIME 2

ATTITUDES	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
INCENTIVES	10.00	3.28	28	9.65	3.47	20	+0.36	+.052
MODERN TECH	13.96	3.20	28	12.50	3.73	20	+1.46	+.210
PARENTS	10.79	1.95	28	10.25	2.51	20	+0.83	+.122
DEDICATION	16.19	3.44	27	15.60	3.95	20	+0.54	+.081

toward Students" and "Attitude toward Relaxed Discipline", most of the observed change arises from the decline in attitudes among the teachers in the CON school. Thus, the effect, if any, is an indirect one for these two indexes. For the third index, "Attitude toward Peer Tutoring", the change in t-values is due primarily to a rise in the mean among the teachers at the EXP school.

Turning now to an examination of the indexes among teachers in Oakland, it is found that very few changes occur, and that some of these are themselves inconsistent. The results for Oakland are presented in Table III-60. In Oakland, the only index drawn from the Teacher Questionnaire which shows a change worth noting is that for "Attitude toward Incentives to Teachers". The t-value for this index at Time 1 is -0.47, and at Time 2 the corresponding t-value is +0.80. This, then, is a shift in the predicted direction, although not a large one. Because this is the only one of the Questionnaire-based indexes which shows a definite shift, and even it changes only a modest amount, it seems fair to say that the incentives model did not create any large changes in teacher attitudes in Oakland.

An examination of the indexes derived from the Teacher Interview for the Oakland data shows that there are two shifts large enough to be noted. One of these is on the index which reflects "Attitude toward Incentives" and the other is on the index for "Relations with Parents". The surprising point about the first of these shifts, the one on "Attitude toward Incentives", is that it is in the direction opposite to the corresponding index from the Teacher Questionnaire. That is, the Interview-based index shows a t-value of +1.73 at Time 1, and a t-value of only +0.36 at Time 2. This shift, then, is opposite in direction to that from the Questionnaire-based index, and also may be the result of regression toward the mean. Inspection of the actual mean scores for each of these indexes at each time point indicates that the shift on the Questionnaire-based index is the result of changes in mean level in both the EXP school and the CON school at Oakland. The mean scores for the Interview-based index indicate that the shift on that index is mainly the result of a decline in the attitude among the teachers in the EXP school. This does not resolve the paradox, but

(Text resumes on page III-149)

TABLE III-60-1

TEACHER QUESTIONNAIRE
E/C, OAKLAND, TIME 1

ATTITUDES	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
ATT TO STDNTS	14.18	2.96	11	13.47	2.10	15	+0.72	+.146
ATT TO SCHOOL	9.20	1.48	10	9.53	1.81	15	-0.48	-.100
INCEN TO TEACH	3.70	1.89	10	4.14	2.54	14	-0.47	-.099
INCEN TO PARNT	3.60	1.90	10	4.38	2.02	13	-0.95	-.202
RELAXED DISC	5.27	1.74	11	6.27	1.49	15	-1.57	-.305
ATT PRIN RE INNOV	7.09	2.07	11	8.00	1.65	15	-1.25	-.247
COOP FACULTY	5.40	1.84	10	6.69	0.87	16	-2.42	-.443
ADULT REL	12.22	3.19	9	14.73	2.05	15	-2.36	-.449
INDIV INST	8.09	1.76	11	8.33	1.91	15	-0.33	-.067
ATT RE PARNT	7.91	1.70	11	7.44	1.26	16	+0.83	+.163
PEER TUTORING	7.92	1.78	12	8.19	2.14	16	-0.36	-.069

TABLE III-60-2

TEACHER QUESTIONNAIRE
E/C, OAKLAND, TIME 2

ATTITUDES	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
ATT TO STDNTS	13.64	2.01	11	13.38	1.66	13	+0.34	+.071
ATT TO SCHOOL	9.42	1.98	12	9.31	1.84	13	+0.14	+.030
INCEN TO TEACH	4.42	2.35	12	3.62	2.63	13	+0.80	+.165
INCEN TO PARNT	4.17	2.12	12	4.38	2.72	13	-0.22	-.046
RELAXED DISC	5.18	2.48	11	6.54	1.61	13	-1.61	-.325
ATT PRIN RE INNOV	7.54	1.61	13	8.08	2.02	13	-0.75	-.152
COOP FACULTY	6.00	1.60	12	7.23	1.09	13	-2.27	-.427
ADULT REL	13.50	2.81	12	15.31	2.21	13	-1.79	-.350
INDIV INST	7.42	1.62	12	8.17	1.90	12	-1.04	-.217
ATT RE PARNT	7.15	1.91	13	7.15	1.68	13	+0.00	+.000
PEER TUTORING	8.38	1.94	13	8.92	1.44	13	-0.80	-.162

TABLE III-60-3

TEACHER INTERVIEW
E/C, OAKLAND
TIME 1

ATTITUDES	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
INCENTIVES	7.76	4.01	17	5.62	3.81	24	+1.73	+.268
MODERN TECH	13.94	4.32	17	12.83	3.57	24	+0.90	+.142
PARENTS	12.12	2.18	17	13.33	1.05	24	-2.38	-.356
DEDICATION	18.59	1.87	17	17.96	1.88	24	+1.06	+.167

TIME 2

ATTITUDES	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
INCENTIVES	5.46	2.57	13	5.09	3.16	23	+0.36	+.062
MODERN TECH	12.79	3.77	14	12.79	3.02	24	-0.01	-.000
PARENTS	11.67	2.16	15	11.87	2.03	24	-0.30	-.050
DEDICATION	17.67	2.92	15	17.22	2.33	23	+0.53	+.087

does indicate that the differences are not entirely due to changes among the EXP teachers. It should be remembered that the difference in the time of the first wave data collection for the two instruments, plus the fact that the response rate to the Teacher Questionnaire was not extremely high, may also be contributing to the lack of agreement in the results.

One final note on these results is that the shift over time that was observed between the EXP and CON schools with respect to the Interview-based index of "Relations with Parents" is due to a mixture of a decline in the level of this index among the teachers in the CON school, and a lesser decline on the same index among the teachers in the EXP school. Thus, this index, like the others, does not provide clear evidence of a positive impact of the incentives model. Instead, it indicates that there may be some benefit of the incentives model in slowing the rate of decline in relations of teachers with parents. Coupled with the other failures of the two instruments to indicate any sizable changes, this lack of consistency and positive impact leads to a conclusion that the incentives model had no general and positive impact in Oakland, as far as the attitudes of the teachers are concerned.

The last of the four individual sites to be considered is San Antonio. The results for San Antonio are presented in Table III-61. They show a somewhat more positive pattern than any of the preceding three. However, these results are not extremely strong.

In San Antonio, two of the indexes from the Teacher Questionnaire show changes from Time 1 to Time 2 large enough to be noted. These are: "Attitude toward Students" and "Attitude toward Relaxed Discipline". Both of these changes are in the predicted direction. The change on the index for "Attitude toward Students", however, could well be due to regression effects. The t-value for this index at Time 1 is -1.14, and the corresponding t-value at Time 2 is -0.05. The pattern on the index for "Attitude toward Relaxed Discipline" is less likely to be the result of a regression effect, since the t-values change from -1.05 and +0.70. It should also be noted that the shift in t-value for the index of "Attitude toward Peer

(Text resumes on page III-153)

TABLE III-61-1

TEACHER QUESTIONNAIRE
E/C, SAN ANTONIO, TIME 1

ATTITUDES	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\bar{X}$	S.D.	N	$\bar{X}$	S.D.	N	t-val.	r_{pb}
ATT TO STDNTS	12.76	3.49	17	14.27	3.32	11	-1.14	-.218
ATT TO SCHOOL	8.18	3.07	17	9.91	1.97	11	-1.66	-.309
INCEN TO TEACH	5.00	2.07	16	4.18	1.40	11	+1.14	+.223
INCEN TO PARNT	4.75	1.77	16	4.45	1.81	11	+0.42	+.084
RELAXED DISC	5.12	1.50	16	5.73	1.42	11	-1.05	-.205
ATT PRIN RE INNOV	8.82	1.24	17	8.45	1.29	11	+0.76	+.147
COOP FACULTY	6.88	2.18	17	7.00	1.70	10	-0.15	-.030
ADULT REL	15.71	2.97	17	15.60	2.72	10	+0.09	+.017
INDIV INST	7.94	1.75	17	8.09	1.64	11	-0.23	-.045
ATT RE PARNT	6.53	1.84	17	7.55	1.81	11	-1.44	-.271
PEER TUTORING	8.06	1.43	17	9.00	1.61	11	-1.62	-.302

TABLE III-61-2

TEACHER QUESTIONNAIRE
E/C, SAN ANTONIO, TIME 2

ATTITUDES	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
ATT TO STDNTS	13.60	3.11	15	13.67	3.46	9	-0.05	-.010
ATT TO SCHOOL	8.27	3.45	15	10.25	1.75	8	-1.51	-.313
INCEN TO TEACH	4.40	1.80	15	3.11	1.96	9	+1.64	+.330
INCEN TO PARNT	4.60	1.99	15	3.56	1.59	9	+1.33	+.274
RELAXED DISC	5.62	1.54	16	5.22	0.97	9	+0.70	+.145
ATT PRIN RE INNOV	9.44	1.09	16	9.00	0.87	9	+1.03	+.210
COOP FACULTY	6.25	2.57	16	7.11	0.93	9	-0.96	-.197
ADULT REL	15.69	3.30	16	16.11	1.27	9	-0.37	-.076
INDIV INST	7.31	1.74	16	8.11	1.62	9	-1.13	-.229
ATT RE PARNT	6.75	2.24	16	7.56	1.67	9	-0.94	-.192
PEER TUTORING	7.60	1.80	15	9.33	1.12	9	-2.59	-.483

TABLE III-61-3

TEACHER INTERVIEW
E/C, SAN ANTONIO
TIME 1

ATTITUDES	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
INCENTIVES	6.67	3.82	18	7.36	3.46	14	-0.53	-.096
MODERN TECH	11.83	2.96	18	12.07	3.50	14	-0.21	-.039
PARENTS	9.28	2.65	18	10.71	1.44	14	-1.82	-.316
DEDICATION	17.00	2.81	18	17.29	1.82	14	-0.33	-.060

TIME 2

ATTITUDES	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
INCENTIVES	8.06	4.30	18	7.08	3.68	13	+0.66	+.122
MODERN TECH	11.00	2.94	17	14.42	2.47	12	-3.29	-.535
PARENTS	10.12	2.32	17	10.69	1.75	13	-0.75	-.014
DEDICATION	15.67	3.11	18	14.08	3.25	13	+1.38	+.248

Tutoring" shows a change that is just on the borderline of being large enough to note. This change, in which the t-value moves from -1.62 at Time 1, to -2.59 at Time 2, is in a negative direction and is not likely to be the result of a regression effect.

The results for the Teacher Interview indexes in San Antonio tend to confirm those from the Questionnaire-based indexes. All four of the Interview-based indexes show a change that is large enough to be noted. For three of the indexes, the changes are in the direction predicted by the hypotheses. The Interview-based index for "Attitude toward Incentives" shows a shift from an initial t-value of -0.53 to a later t-value of +0.66. This change, while not very large, is in the direction predicted by the hypothesis. The index for "Attitude toward Modern Techniques" in San Antonio is the one which changes in the direction opposite to the prediction made by the hypothesis. This index shows a t-value of -0.21 at Time 1, and a t-value of -3.29 at Time 2. This change is fairly large, and is unlikely to be due to any regression effects, since it indicates a clear divergence of the means of the two schools over time.

The Interview-based indexes for "Relations with Parents" and for "Dedication to Teaching" also show changes in the predicted direction. The change observed on the "Relations with Parents" index (from a t-value of -1.82 to a t-value of -0.75) might be due to regression toward the mean, but this explanation cannot be offered for the shift observed in the "Dedication to Teaching" index. For that index, the Time 1 t-value is -0.33 and the Time 2 t-value is +1.38.

The results in San Antonio tend to suggest a situation in which the participants generally became more positive in their actions over time, but did not adopt the kinds of teaching techniques usually thought of as being new and modern. This impression is confirmed by an examination of the actual means for the various indexes. For "Attitude toward Peer Tutoring" and for "Attitude toward Individualized Instruction", there is a decline in the average score obtained in the EXP school from Time 1 to Time 2. Although these declines are not large enough to create substantial

changes in the t-values, they are consistent with the results from the Interview-based index of "Attitude toward Modern Techniques".

The overall conclusions that can be drawn from these comparisons of Teacher Questionnaire indexes and Teacher Interview indexes at the four separate sites must be tentative. However, it would appear that the incentives model did not produce any positive impacts in Cincinnati, and may indeed have produced a negative impact. In light of the other known information about Cincinnati, this should perhaps be stated instead as reflecting the disorganization observed in the EXP school because of the prolonged absence of the principal, and the competitive striving by the CON school in Cincinnati to demonstrate that it could excel without being awarded any incentives. The impacts in Jacksonville, though scattered, seem generally to be in a positive direction. The impacts of the incentives model in Oakland, on the other hand, are inconsistent, but probably negative. In any event, there is no evidence in Oakland for the contention that the incentives model had a positive effect. Finally, the situation in San Antonio seems to be that there was a general beneficial impact of the treatment, but no increase in use of modern techniques. On the contrary, the results suggest that in the San Antonio EXP school, the use of modern techniques was lessened during the project.

Model-Based Comparisons--Teacher Questionnaire and Interview Results

In view of the generally erratic results from the individual site comparisons, the model-based comparisons will not be discussed in detail. To present such discussion would implicitly suggest that the combined results are directly interpretable. Given the lack of consistency between the two sites which experienced the same incentives model (a phenomenon also noted in the analysis of the achievement data, and the student attitude and behavior data), the model-based comparisons are of questionable meaning. However, for the benefit of the interested reader, Tables III-62 through III-65 are presented showing these results.

The comparisons for the two sites at which the Teacher-Only model was used (Cincinnati and Jacksonville) indicate a slight shift in favor of the EXP school in the Questionnaire-based indexes for "Attitude toward Relaxed Discipline" and "Attitude toward Peer Tutoring". These results are presented in Table III-62. However, the index of "Attitude toward Modern Techniques" from the Teacher Interview data shows a shift in the direction contrary to the hypothesis, although that shift is just slightly too small to be noted.

In the comparisons involving only the schools at the Parent-Teacher model sites (Oakland and San Antonio), there is a mixed pattern of results. These results are presented in Table III-63. There are shifts in favor of the EXP schools on the Questionnaire-based indexes of "Attitude toward Incentives to Teachers" and "Attitude toward Relaxed Discipline". However, there are shifts in the direction opposite to that predicted on the Questionnaire-based indexes of "Perception of Cooperation among Faculty", "Attitude toward Individualized Instruction", and "Attitude toward Peer Tutoring". Consistent with the negative patterns just mentioned, the index for "Attitude toward Modern Techniques" based on the Teacher Interview also shows a shift in the direction contrary to that hypothesized. Thus, it would appear that in the PT sites, there was some general tendency to decrease the use of modern techniques in the EXP schools, as compared to the level of their use in the CON schools. The results in the PT sites from the Teacher Interview also indicate a relative improvement in the index for "Relations with Parents".

As mentioned in the earlier sections where similar comparisons were made, the PT-E versus TO-E comparison is one in which site-to-site differences can be confounded with any actual treatment impacts, and so it must be interpreted with extra caution. However, it does provide some information as to the possible level of the marginal impact of the Parent-Teacher model over and above the impact of the Teacher-Only model. The results of this comparison are given in Table III-64. Only two of the indexes from the Teacher Questionnaire show differences over time in this set of comparisons. They are the index for "Perception of the Principal's Attitude

(Text resumes on page III-165)

TABLE III-62-1

TEACHER QUESTIONNAIRE
E/C, TEACHER ONLY, TIME 1

ATTITUDES	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
ATT TO STDNTS	13.11	2.91	27	11.63	3.93	19	+1.47	+.216
ATT TO SCHOOL	9.42	3.34	26	8.89	2.98	19	+0.55	+.083
INCEN TO TEACH	6.04	2.53	28	5.42	2.52	19	+0.82	+.121
INCEN TO PARNT	6.00	2.45	28	4.89	2.71	19	+1.46	+.212
RELAXED DISC	6.29	1.67	28	6.11	2.26	19	+0.31	+.047
ATT PRIN RE INNOV	7.64	2.92	28	8.32	1.49	19	-0.92	-.136
COOP FACULTY	7.30	2.25	27	6.79	2.23	19	+0.76	+.113
ADULT REL	15.15	4.71	27	15.11	3.45	19	+0.03	+.000
INDIV INST	8.89	1.45	28	8.16	1.34	19	+1.76	+.253
ATT RE PARNT	6.93	1.54	27	6.16	1.86	19	+1.53	+.224
PEER TUTORING	8.93	1.44	28	8.47	1.78	19	+0.97	+.143

TABLE III-62-2

TEACHER QUESTIONNAIRE
E/C, TEACHER ONLY, TIME 2

ATTITUDES	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
ATT TO STDNTS	14.00	3.28	19	11.00	4.99	8	+1.86	+.348
ATT TO SCHOOL	9.00	3.76	20	7.75	3.49	8	+0.81	+.157
INCEN TO TEACH	5.48	2.27	21	4.67	2.24	9	+0.90	+.167
INCEN TO PARNT	5.70	2.49	20	4.22	2.39	9	+1.49	+.276
RELAXED DISC	6.30	1.66	20	5.33	1.94	9	+1.38	+.257
ATT PRIN RE INNOV	6.81	3.06	21	8.11	1.36	9	-1.22	-.224
COOP FACULTY	6.62	2.54	21	6.33	2.50	9	+0.28	+.054
ADULT REL	13.43	5.40	21	14.44	3.50	9	-0.52	-.097
INDIV INST	8.75	1.74	20	7.67	1.50	9	+1.61	+.296
ATT RE PARNT	6.79	1.30	21	6.22	1.92	9	+0.90	+.168
PEER TUTORING	9.40	0.94	20	8.33	1.41	9	+2.41	+.421

TABLE III-62-3

TEACHER INTERVIEW
E/C, TEACHER ONLY
TIME 1

ATTITUDES	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
INCENTIVES	10.07	3.18	40	8.65	3.24	37	+1.95	+.219
MODERN TECH	12.65	3.22	40	10.73	3.15	37	+2.64	+.292
PARENTS	11.42	2.24	40	10.92	2.17	37	+1.01	+.115
DEDICATION	17.72	2.10	40	17.30	2.96	37	+0.54	+.085

TIME 2

ATTITUDES	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
INCENTIVES	9.49	3.16	41	8.26	3.48	35	+1.62	+.185
MODERN TECH	13.46	3.12	41	12.14	3.73	36	+1.70	+.192
PARENTS	10.44	2.06	41	10.29	2.35	35	+0.30	+.035
DEDICATION	16.77	3.11	40	16.58	3.34	36	+0.26	+.030

TABLE III-63-1

TEACHER QUESTIONNAIRE
E/C, PARENT TEACHER ONLY, TIME 1

ATTITUDES	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
ATT TO STDNTS	13.32	3.31	28	13.81	2.65	26	-0.59	-.082
ATT TO SCHOOL	8.56	2.61	27	9.69	1.85	26	-1.82	-.248
INCEN TO TEACH	4.50	2.06	26	4.16	2.08	25	+0.59	+.084
INCEN TO PARNT	4.31	1.87	26	4.42	1.89	24	-0.20	-.030
RELAXED DISC	5.19	1.57	27	6.04	1.46	26	-2.05	-.276
ATT PRIN RE INNOV	8.14	1.80	28	8.19	1.50	26	-0.11	-.014
COOP FACULTY	6.33	2.15	27	6.81	1.23	26	-0.98	-.136
ADULT REL	14.50	3.43	26	15.08	2.33	25	-0.70	-.100
INDIV INST	8.00	1.72	28	8.23	1.77	26	-0.49	-.067
ATT RE PARNT	7.07	1.88	28	7.48	1.48	27	-0.90	-.122
PEER TUTORING	8.00	1.56	29	8.52	1.95	27	-1.10	-.148

TABLE III-63-2

TEACHER QUESTIONNAIRE
E/C, PARENT TEACHER ONLY, TIME 2

ATTITUDES	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
ATT TO STDNTS	13.62	2.65	26	13.50	2.48	22	+0.15	+.022
ATT TO SCHOOL	8.78	2.90	27	9.67	1.83	21	-1.23	-.178
INCEN TO TEACH	4.41	2.02	27	3.41	2.34	22	+1.60	+.227
INCEN TO PARNT	4.41	2.02	27	4.05	2.32	22	+0.58	+.085
RELAXED DISC	5.44	1.95	27	6.00	1.51	22	-1.09	-.158
ATT PRIN RE INNOV	8.59	1.64	29	8.45	1.68	22	+0.28	+.040
COOP FACULTY	6.14	2.17	28	7.18	1.01	22	-2.07	-.287
ADULT REL	14.75	3.24	28	15.64	1.89	22	-1.14	-.162
INDIV INST	7.36	1.66	28	8.14	1.74	21	-1.61	-.228
ATT RE PARNT	6.93	2.07	29	7.32	1.64	22	-0.72	-.102
PEER TUTORING	7.96	1.88	28	9.09	1.31	22	-2.40	-.327

TABLE III-63-3

TEACHER INTERVIEW
E/C, PARENT TEACHER ONLY
TIME 1

ATTITUDES	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
INCENTIVES	7.20	3.89	35	6.26	3.73	38	+1.05	+.124
MODERN TECH	12.86	3.78	35	12.55	3.52	38	+0.36	+.042
PARENTS	10.66	2.80	35	12.37	1.75	38	-3.16	-.351
DEDICATION	17.77	2.50	35	17.71	1.86	38	+0.12	+.014

TIME 2

ATTITUDES	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
INCENTIVES	6.97	3.85	31	5.81	3.45	36	+1.30	+.160
MODERN TECH	11.81	3.40	31	13.33	2.92	36	-1.98	-.238
PARENTS	10.84	2.34	32	11.46	1.99	37	-1.18	-.142
DEDICATION	16.58	3.14	33	16.08	3.06	36	+0.66	+.080

TABLE III-64-1

TEACHER QUESTIONNAIRE
PARENT TEACHER-E/TEACHER ONLY, TIME 1

ATTITUDES	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
ATT TO STDNTS	13.32	3.31	28	13.11	2.91	27	+0.25	+.035
ATT TO SCHOOL	8.56	2.61	27	9.42	3.34	26	-1.06	-.146
INCEN TO TEACH	4.50	2.06	26	6.04	2.53	28	-2.43	-.320
INCEN TO PARNT	4.31	1.87	26	6.00	2.45	28	-2.84	-.366
RELAXED DISC	5.19	1.57	27	6.29	1.67	28	-2.51	-.326
ATT PRIN RE INNOV	8.14	1.80	28	7.64	2.92	28	+0.77	+.104
COOP FACULTY	6.33	2.15	27	7.30	2.25	27	-1.61	-.218
ADULT REL	14.50	3.43	26	15.15	4.71	27	-0.57	-.079
INDIV INST	8.00	1.72	28	8.89	1.45	28	-2.10	-.275
ATT RE PARNT	7.07	1.83	28	6.93	1.54	27	+0.31	+.042
PEER TUTORING	8.00	1.56	29	8.93	1.44	28	-2.34	-.300

TABLE III-64-2

TEACHER QUESTIONNAIRE
PARENT TEACHER-E/TEACHER ONLY-E, TIME 2

ATTITUDES	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
ATT TO STDNTS	13.62	2.65	26	14.00	3.28	19	-0.43	-.066
ATT TO SCHOOL	8.78	2.90	27	9.00	3.76	20	-0.23	-.035
INCEN TO TEACH	4.41	2.02	27	5.48	2.27	21	-1.72	-.246
INCEN TO PARNT	4.41	2.02	27	5.70	2.49	20	-1.96	-.281
RELAXED DISC	5.44	1.95	27	6.30	1.66	20	-1.58	-.230
ATT PRIN RE INNOV	8.59	1.64	29	6.81	3.06	21	+2.65	+.358
COOP FACULTY	6.14	2.17	28	6.62	2.54	21	-0.71	-.102
ADULT REL	14.75	3.24	28	13.43	5.40	21	+1.07	+.154
INDIV INST	7.36	1.66	28	8.75	1.74	20	-2.81	-.382
ATT RE PARNT	6.93	2.07	29	6.79	1.30	21	+0.33	+.048
PEER TUTORING	7.96	1.88	28	9.40	0.94	20	-3.15	-.421

TABLE III-64-3

TEACHER INTERVIEW
PARENT TEACHER-E/TEACHER ONLY-E
TIME 1

ATTITUDES	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
INCENTIVES	7.20	3.89	35	10.07	3.18	40	-3.52	-.381
MODERN TECH	12.86	3.78	35	12.65	3.22	40	+0.26	+.030
PARENTS	10.66	2.80	35	11.42	2.24	40	-1.32	-.153
DEDICATION	17.77	2.50	35	17.72	2.50	35	+0.09	+.010

TIME 2

ATTITUDES	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
INCENTIVES	6.97	3.85	31	9.49	3.16	41	-3.05	-.343
MODERN TECH	11.81	3.40	31	13.46	3.12	41	-2.15	-.248
PARENTS	10.84	2.34	32	10.44	2.06	41	+0.78	+.093
DEDICATION	16.58	3.14	33	16.77	3.11	40	-0.27	-.032

toward Innovation" and the "Adult Relations" index. The latter of these is, as indicated in Table III-56, a composite index which has the index for "Perception of Principal's Attitude toward Innovation" as one of its components. Therefore, these two results actually are reflecting a single phenomenon, namely that the teachers at the experimental schools in the PT sites find their principal more receptive than do the teachers in the experimental schools at the TO sites. In view of the replacement of one of the principals in the experimental school in a TO site (Cincinnati) for reasons of ill health, this result is not surprising. The Teacher Interview indexes for the comparison of PT-E versus TO-E indicate, as did the same indexes in the preceding comparison of schools at the PT sites, that there is a relative decline in "Attitude toward Modern Techniques" in the Parent-Teacher experimental schools as compared with the Teacher-Only experimental schools. Also, this set of comparisons indicates that the PT-E schools shift to a relatively more favorable position than the TO-E schools as far as "Relations with Parents" are concerned. But it must be remembered that this is due primarily to a slower rate of decline in the EXP schools.

The last of the model-based comparisons is the overall combination of all four experimental schools compared with all four control schools. These results are presented in Table III-65. From this set of comparisons, only one of the 11 indexes based on the Questionnaire shows a sizable shift. This is the index for "Perception of Cooperation among Faculty" which shifts in the direction opposite to that predicted by the hypothesis. Of the four indexes from the Teacher Questionnaire, there is likewise only one which shows a shift large enough to be noted. This is the index for "Attitude toward Modern Techniques" which also shifts in the direction opposite the hypothesis.

<u>Teacher Questionnaire and Interview Results--
Discussion</u>

The preceding pages have presented the results obtained from the Teacher Questionnaire and Teacher Interview data. As far as any confirmation of the hypothesized relationships is concerned, there is relatively

(Text resumes on page III-169)

TABLE III-65-1

TEACHER QUESTIONNAIRE
E/C, ALL, TIME 1

ATTITUDES	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
ATT TO STDNTS	13.22	3.10	55	12.89	3.39	45	+0.51	+0.51
ATT TO SCHOOL	8.98	2.99	53	9.36	2.39	45	-0.68	-.069
INCEN TO TEACH	5.30	2.42	54	4.70	2.34	44	+1.22	+.124
INCEN TO PARNT	5.19	2.33	54	4.63	2.27	43	+1.18	+.120
RELAXED DISC	5.75	1.70	55	6.07	1.81	45	-0.91	-.092
ATT PRIN RE INNOV	7.89	2.42	56	8.24	1.48	45	-0.85	-.085
COOP FACULTY	6.81	2.23	54	6.80	1.70	45	+0.04	+.000
ADULT REL	14.83	4.11	53	15.09	2.83	44	-0.36	-.036
INDIV INST	8.45	1.64	56	8.20	1.59	45	+0.76	+.076
ATT RE PARNT	7.00	1.71	55	6.93	1.76	46	+0.19	+.020
PEER TUTORING	8.46	1.56	57	8.50	1.86	46	-0.13	-.014

TABLE III-65-2

TEACHER QUESTIONNAIRE
E/C, ALL, TIME 2

ATTITUDES	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\bar{X}$	S.D.	N	$\bar{X}$	S.D.	N	t-val.	r_{pb}
ATT TO STDNTS	12.78	2.91	45	12.83	3.42	30	+1.28	+.149
ATT TO SCHOOL	8.87	3.25	47	9.14	2.49	29	-0.38	-.044
INCEN TO TEACH	4.87	2.18	48	3.77	2.35	31	+2.13	+.236
INCEN TO PARNT	4.96	2.30	47	4.10	2.30	31	+1.62	+.182
RELAXED DISC	5.81	1.86	47	5.81	1.64	31	+0.00	+.000
ATT PRIN RE INNOV	7.84	2.48	50	8.35	1.58	31	-1.03	-.115
COOP FACULTY	6.35	2.32	49	6.94	1.59	31	-1.24	-.139
ADULT REL	14.18	4.30	49	15.29	2.47	31	-1.30	-.146
INDIV INST	7.94	1.81	48	8.00	1.66	30	-0.15	-.017
ATT RE PARNT	6.86	1.77	50	7.00	1.77	31	-0.35	-.039
PEER TUTORING	8.56	1.70	48	8.87	1.36	31	-0.85	-.096

TABLE III-65-3

TEACHER INTERVIEW
E/C, ALL
TIME 1

ATTITUDES	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
INCENTIVES	8.73	3.79	75	7.44	3.68	75	+2.12	+.172
MODERN TECH	12.75	3.47	75	11.65	3.44	75	+1.94	+.157
PARENTS	11.07	2.53	75	11.66	2.08	75	-1.55	-.126
DEDICATION	17.75	2.28	75	17.51	2.46	75	+0.62	+.051

TIME 2

ATTITUDES	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
INCENTIVES	8.40	3.67	72	7.01	3.65	71	+2.27	+.188
MODERN TECH	12.75	3.33	72	12.74	3.38	72	+0.02	+.000
PARENTS	10.62	2.18	73	10.89	2.24	72	-0.74	-.062
DEDICATION	16.68	3.10	73	16.33	3.19	72	+0.67	+.057

little support in these data. The positive outcomes are few in number and are scattered throughout the sites and indexes. For the individual sites, of 44 hypotheses concerned with Teacher Questionnaire indexes, only six hypotheses turn out to be notable and in the predicted direction. At the same time, there are two hypotheses from the Teacher Questionnaire for which the outcomes were large enough to be notable, but in a direction opposite to that predicted. Similarly, of 16 predictions made for the Teacher Interview indexes, there were only four positive confirmations, matched by four instances in which the results were in the direction opposite to expectations.

Despite the sparseness of these results, and the occurrence of outcomes opposite to those hypothesized, there are some general tendencies that may well reflect impacts of the incentives model. These have been mentioned in connection with the individual sites, and will be briefly reviewed here.

One tendency that is contrary to the direction predicted is the relative decrease in favorable attitudes toward modern approaches to teaching that was observed in the EXP schools of Cincinnati and San Antonio. This outcome perhaps indicates that one reaction to the introduction of the incentives model is a return to the traditional approaches. It may be that teachers feel these "tried-and-true" approaches will be as effective as any for the purpose of producing measurable increases in Reading and Math achievement. A second instance of a possible general pattern is in the relative increases in the Parent-Teacher EXP schools in the relations between teachers and parents, reported by the teachers on the interview. This result may be due, of course, to causes other than the PT model. It is created, in part, not by an increase in favor of the EXP schools, but simply by a lesser rate of decline. Thus, it should be regarded as tentative.

Two general points can be made about these results. First, the findings opposite in direction to the hypothesis provide an example of one limitation that is inherent in this stage of research. Because it is relatively easy to explain the shift toward less favorable attitudes to modern techniques after the data have been presented, a critic might argue that it

would have been more desirable to have anticipated this possibility at the time the data collection was being planned, and to have included additional efforts so that relevant information establishing the nature of this process would be available.

The difficulty with this suggestion as a general strategy, of course, is that there are a large number of alternate explanations which can easily be proposed, and an even larger number of possible data patterns which might be found. Thus, to attempt to cover all of these in a work aimed at preliminary knowledge would not be possible. Rather, it is the function of a first study such as this to narrow down somewhat the possibilities, so that subsequent research on the topic can in turn be concentrated more effectively.

The second general point is that these results, like those reported in earlier sections, may reflect largely or entirely the impact of what James Coleman has called "random shocks" (Coleman, 1964). This is not to say that there are not systematic processes occurring, but that they are made less detectable by accidental fluctuations in the level of the outcome variables. These "random shocks" act much like static in a radio, and make the reception of the message being transmitted difficult or impossible. In this project, the message being transmitted is whatever is really the impact of the treatment at the various sites. An understanding of that message is lessened, however, by the presence of additional components. The message is also made less understandable by whatever systematic distortion may be introduced by the instruments--in this case, by the techniques of data collection employed. The amount of this distortion is not precisely known, and so an additional measure of uncertainty is present.

The general remedy for a situation in which communication is faulty is to use whatever other channels and media may be available. In the present study, this approach was adopted in the form of collecting a variety of data from several sources. The next section takes up the analysis of one of the additional sources of data, the Teacher Classroom Observation data.

Teacher Classroom Observation Results--Introduction

Teacher Classroom Observations were conducted by the TAC field representatives at the same time and in the same manner as the Student Classroom Observations were made. Table III-66 presents the 14 teacher activities and the four context categories, along with the direction of difference hypothesized for each of the 18 variables.

Single-Site Comparisons--Teacher Classroom Observation Results

Tables III-67 through III-74 present the results for the four sites. As in the analysis of the Student Classroom Observation results, formal tests of significance have not been used in the analysis of the Teacher Classroom Observation results. Rather, the analysis consisted of a search for patterns of change by means of comparing the E/C t-value at Time 1 with the corresponding t-value at Time 2. Although this method of analysis may be unappealing to readers who are interested in short, simple answers, it is believed that the method is in keeping with the complex nature of the project. That is, this was a field project using only a few sites spread across the country. No attempt was made in the design to randomize for even the known sources of variance. Furthermore, there were delays in the implementation of the project and varying levels of knowledge among the participants at the four sites about the project objectives. In short, to present a few simple results in the analysis would be to imply that simple questions had been asked. This most clearly was not the intent or the fact in this project.

Tables III-67 and III-68 present the Teacher Observation results for Cincinnati. From a comparison of the Time 1 and Time 2 t-values, there were four teacher activity changes in favor of the EXP school as predicted. In Item 1, "Testing"; Item 6, "Praising"; Item 12, "Supervising", the t-value differences from Time 1 to Time 2 were all greater than +1.99. Less substantial changes in favor of the EXP school were found in Item 2, "Explaining" and Item 12, "Supervising" with t-value differences of +1.07 and +1.99, respectively.

(Text resumes on page III-175)

TABLE III-66

TEACHER CLASSROOM OBSERVATION

CONTEXT/ ITEM	EXPLANATION	HYP
1	Testing	E > C
2	Explaining subject	E > C
3	Lecturing Subject	E < C
4	Giving reward: Giving or promising extrinsic motivator	E > C
5	Punishing: Giving or promising physical punishment or removal of privilege	E < C
6	Praising: Student behavior or academic work	E > C
7	Criticizing: Student behavior or academic work	E < C
8	Listening: To subject matter discussion, presentation	E > C
9	Questioning: About subject matter	E > C
10	Drilling: Rote, repetition of formulas, etc.	E < C
11	Administering: Lunch tickets attendance, etc.	E < C
12	Supervising: Of student-initiated work	E > C
13	Peer Tutoring: Establishing, supervising, encouraging	E > C
14	Enriching: Additional, not-for-credit, materials	E > C
One Student		E > C
Small Group	Two-five students	E > C
Large Group	Six-3/4 class	E < C
Whole Class		E < C

TABLE III-67

TEACHER CLASSROOM OBSERVATION
E/C, CINCINNATI, TIME 1

ITEMS & CONTEXT	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
ITEM 1	0.52	3.74	56	1.28	5.42	64	-0.89	-.081
ITEM 2	9.23	6.20	56	9.89	6.34	64	-0.66	-.060
ITEM 3	1.18	4.77	56	0.75	2.62	64	+0.62	+.057
ITEM 4	0.00	0.00	56	0.00	0.00	64	UNDEF	UNDEF
ITEM 5	0.02	0.13	56	0.02	0.12	64	+0.09	+.010
ITEM 6	0.91	1.64	56	0.83	1.56	64	+0.28	+.026
ITEM 7	1.98	7.51	56	1.42	2.16	64	+1.25	+.114
ITEM 8	5.66	5.23	56	5.08	4.32	64	+0.67	+.062
ITEM 9	5.66	4.20	56	7.95	5.82	64	-2.44	-.219
ITEM 10	1.73	4.09	56	0.70	2.04	64	+1.78	+.162
ITEM 11	8.12	8.31	56	5.89	6.85	64	+1.61	+.147
ITEM 12	0.41	2.21	56	0.05	0.28	64	+1.31	+.120
ITEM 13	0.04	0.27	56	0.00	0.00	64	UNDEF	UNDEF
ITEM 14	0.52	2.52	56	0.39	2.88	64	+0.26	+.024
ONE STUD	5.80	6.56	56	5.03	7.79	64	+0.58	+.054
SM GROUP	4.12	7.53	56	2.64	7.97	64	+1.04	+.096
LG GROUP	5.36	10.36	56	10.67	13.41	64	-2.40	-.216
WHOL CLAS	20.43	13.17	56	15.98	15.60	64	+1.67	+.152

TABLE III-68

TEACHER CLASSROOM-OBSERVATION
E/C, CINCINNATI, TIME 2

ITEMS & CONTEXT	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\bar{X}$	S.D.	N	$\bar{X}$	S.D.	N	t-val.	r_{pb}
ITEM 1	1.15	5.52	72	0.06	0.56	80	+1.76	+.142
ITEM 2	7.03	6.88	72	6.57	6.47	80	+0.42	+.035
ITEM 3	0.87	2.79	72	0.79	4.19	80	+0.15	+.010
ITEM 4	0.03	0.24	72	0.00	0.00	80	+1.05	+.086
ITEM 5	0.04	0.20	72	0.01	0.11	80	+1.12	+.091
ITEM 6	1.01	1.81	72	0.37	1.04	80	+2.70	+.215
ITEM 7	1.71	2.43	72	0.87	1.59	80	+2.53	+.202
ITEM 8	4.03	5.65	72	3.41	4.67	80	+0.73	+.060
ITEM 9	4.85	5.39	72	6.45	7.48	80	-1.50	-.122
ITEM 10	0.79	2.27	72	0.52	1.46	80	+0.87	+.071
ITEM 11	4.12	7.25	72	3.21	5.05	80	+0.91	+.074
ITEM 12	0.82	2.19	72	0.01	0.11	80	+3.30	+.260
ITEM 13	0.00	0.00	72	0.00	0.00	80	UNDEF	UNDEF
ITEM 14	0.00	0.00	72	0.09	0.60	80	UNDEF	UNDEF
ONE STUD	8.78	8.67	72	4.04	5.67	80	+4.03	+.312
SM GROUP	3.58	7.56	72	3.29	7.56	80	+0.24	+.020
LG GROUP	4.14	8.88	72	7.19	11.28	80	-1.84	-.148
WHOL CLAS	10.17	12.86	72	7.79	11.76	80	+1.19	+.097

On the other hand, there were two teacher activity variables for which the change was opposite to the direction predicted. Although none of these changes is large, the EXP school was above the CON school in Item 5, "Punishing" and Item 7, "Criticizing".

Among the context categories, it appears that the EXP teachers did spend significantly more time with individual students toward the conclusion of the project. The t-value for this context changed from +4.03. This change is due to a substantial increase in the "One Student" context in the EXP school combined with a slight decrease in this instructional context in the CON school.

Overall, the pattern of change is difficult to interpret because in some cases the differences are due primarily to EXP school changes and in other cases to CON school changes, with little change in the EXP school. For example, in Item 2, "Explaining",and Item 6, "Praising", the mean frequency of the CON school decreased a fair amount whereas the EXP school frequency was relatively constant from Time 1 to Time 2. Among the three activities which changed in the direction opposite to that hypothesized, two of the changes, Item 5, "Punishing", and Item 7, "Criticizing", were primarily due to decreases in the observed frequency at the CON school. Taken at face value, the tables would indicate that the EXP school teachers were leading a somewhat schizoid life in the classroom, *i.e.*, combining the more "traditional" behaviors of testing, punishing and criticizing with the "innovations" of individualized instruction, praising and supervising work initiated by the students. That persons, including teachers, can simultaneously govern their activities by several diverse theories of human behavior and motivation is not denied in this analysis. Indeed, evidence for it emerges in the results. However, because the apparent pattern at the EXP school depends also upon what happens in the CON school, this set of results should be regarded with some suspicion. In Cincinnati, the EXP school did increase instruction with individual students, more than doubling the frequency of activity in this context from Time 1 to Time 2. That much is clear and is also supported by the Student Observation results. However, conclusions

beyond this are fraught with danger. Even on Time 1, "Testing", where the EXP school tripled its efforts from Time 1 to Time 2, there is no supporting evidence to be found in the Student Observation data.

Tables III-69 and III-70 present the Teacher Classroom Observation results for Jacksonville. There were nine changes in teacher activity, two of which were in the direction opposite to what was hypothesized. Although the absolute and relative changes were very slight. EXP school teachers less frequently enriched their classrooms and more frequently lectured, neither of which was predicted. Among the seven items in which the direction of change was as predicted, five of the changes were substantial, involving t-value differences from Time 1 to Time 2 greater than +1.99. These substantial changes were found in Item 2, "Explaining"; Item 6, "Praising"; Item 8, "Listening"; Item 9, "Questioning"; and Item 11, "Administrating". Smaller t-value differences from Time 1 to Time 2 were found in Item 1, "Testing" and Item 7, "Criticizing" with the differences being +1.35 and -1.01, respectively.

An examination of the t-values for the four context categories reveals two changes. Consistent with the predicted change, the EXP school improved its status in the use of "Small Group" instruction, with the t-values going from -1.00 at Time 1 to +0.29 at Time 2. Contrary to this slight change, the t-value for "Large Group" instruction went from -0.17 to +3.87, a rather clear change not in the direction hypothesized.

Superficially, the pattern of change in the Jacksonville classrooms appears to be rather clear. That is, eight changes were as hypothesized. However, in two of these variables, "Testing" and "Praising", the major source of the change was the CON school, which reduced the frequency of these activities from Time 1 to Time 2. In four activities, "Explaining", "Criticizing", "Listening", and "Administration", both the EXP and CON schools reduced the frequency of the activity; but the EXP school reduction was less. In the remaining two of the eight, "Questioning" and "Small Group", an increase in the EXP school frequency was matched by a decrease in the CON school frequency, thereby somewhat distorting the magnitude of the change. Furthermore, these latter changes could well be due to regression effect.

TABLE III-69

TEACHER CLASSROOM OBSERVATION
E/C, JACKSONVILLE, TIME 1

ITEMS & CONTEXT	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\bar{X}$	S.D.	N	$\bar{X}$	S.D.	N	t-val.	r_{pb}
ITEM 1	0.62	2.64	112	1.10	3.34	81	-1.10	-.079
ITEM 2	7.72	4.99	112	8.65	5.36	81	-1.24	-.089
ITEM 3	1.12	3.00	112	1.44	3.61	81	-0.67	-.048
ITEM 4	0.05	0.26	112	0.04	0.19	81	+0.48	+.035
ITEM 5	0.09	0.37	112	0.12	0.90	81	-0.36	-.026
ITEM 6	1.81	2.69	112	1.42	2.11	81	+1.09	+.079
ITEM 7	2.92	3.27	112	2.38	3.12	81	+1.15	+.082
ITEM 8	4.91	4.62	112	5.64	4.04	81	-1.14	-.082
ITEM 9	6.89	4.85	112	7.01	4.55	81	-0.17	-.014
ITEM 10	2.78	4.20	112	2.37	4.01	81	+0.68	+.049
ITEM 11	2.25	4.06	112	1.22	2.35	81	+2.04	+.146
ITEM 12	3.91	5.38	112	4.23	6.21	81	-0.40	-.028
ITEM 13	0.16	1.25	112	0.22	1.29	81	-0.33	-.024
ITEM 14	0.51	2.92	112	0.10	0.46	81	+1.25	+.090
ONE STUD	11.89	10.24	112	10.47	8.30	81	+1.03	+.074
SM GROUP	2.53	6.43	112	3.57	8.01	81	-1.00	-.072
LG GROUP	6.65	11.14	112	6.91	10.28	81	-0.17	-.010
WHOL CLAS	14.61	12.39	112	15.05	12.98	81	-0.24	-.017

TABLE III-70

TEACHER CLASSROOM OBSERVATION
E/C, JACKSONVILLE, TIME 2

ITEMS & CONTEXT	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
ITEM 1	0.59	2.40	112	0.51	2.24	100	+0.25	+.017
ITEM 2	8.10	4.26	112	6.50	5.30	100	+2.43	+.166
ITEM 3	2.24	3.39	112	1.72	3.62	100	+1.04	+.071
ITEM 4	0.02	0.13	112	0.02	0.20	100	-0.09	-.000
ITEM 5	0.06	0.34	112	0.13	0.56	100	-1.07	-.074
ITEM 6	1.56	2.01	112	0.74	1.57	100	+3.29	+.222
ITEM 7	2.10	3.05	112	2.04	3.23	100	+0.13	+.010
ITEM 8	5.85	4.75	112	4.70	4.71	100	+1.76	+.121
ITEM 9	7.82	4.89	112	6.00	5.05	100	+2.67	+.181
ITEM 10	2.05	3.48	112	1.87	3.40	100	+0.39	+.026
ITEM 11	1.29	2.45	112	1.36	3.44	100	-0.18	-.014
ITEM 12	4.04	5.42	112	3.94	5.87	100	+0.13	+.010
ITEM 13	0.19	0.88	112	0.48	2.25	100	-1.27	-.088
ITEM 14	0.13	0.81	112	0.23	1.35	100	-0.64	-.044
ONE STUD	8.95	9.57	112	7.65	9.19	100	+1.00	+.069
SM GROUP	2.68	7.42	112	2.40	6.42	100	+0.29	+.020
LG GROUP	9.81	12.85	112	3.87	8.87	100	+3.87	+.258
WHOL CLAS	14.47	13.06	112	16.32	13.83	100	-1.00	-.069

Even in the areas that changed in the direction not predicted, the results are ambiguous. In Item 14, "Enriching", a slight EXP school decrease in this activity was magnified by a slight CON school increase. The apparent EXP school increase in "Large Group" instruction is really the result of the combinations of a small increase there and with a nearly one-half reduction in this context in the CON school. The only clear result, and one that was not predicted, is that the EXP school increase in lecturing is due to a doubling of its frequency of this activity, with little change in the CON school. Unfortunately, this clear change is not matched by an increase in EXP school student "Listening".

The Teacher Classroom Observation results for Oakland are presented in Tables III-71 and III-72. Among the 14 teacher activities, there are discernible trends between the Time 1 and Time 2 t-values in seven of the variables. Five of these trends, however, are in the direction opposite to what was predicted. The Time 1 to Time 2 differences are substantial for Item 8, "Listening", for which the t-value went from +1.72 to -0.36; Item 11, "Administrating", for which the t-value went from -1.39 to +1.19; and for Item 12, "Supervising", for which the t-value went from +1.72 to -0.58. For the remaining two teacher activities, the observed trends are not as substantial: Item 5, "Punishing", had a Time 1 to Time 2 t-value difference of +1.01; and Item 13, "Peer Tutoring", had a Time 1 to Time 2 t-value difference of -1.60.

For the two activities in which the change was in the direction predicted, the results differed in magnitude. In Item 7, "Criticizing", there was a substantial t-value change from +1.89 at Time 1 to -0.58 at Time 2. However, in Item 14, "Enriching", the change was not as pronounced, the t-value changing from -0.52 to +0.55.

There are three trends found in the context categories, one of which was not in the predicted direction. In the "Small Group" context, the t-value went from +1.57 at Time 1 to -1.03 at Time 2, an indication that the frequency of this type of instruction decreased in the EXP school relative to the CON school. However, the frequency of "One Student"

TABLE III-71

TEACHER CLASSROOM OBSERVATION
E/C, OAKLAND, TIME 1

ITEMS & CONTEXT	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\bar{X}$	S.D.	N	$\bar{X}$	S.D.	N	t-val.	r_{pb}
ITEM 1	0.47	2.31	72	0.82	3.29	95	-0.77	-.060
ITEM 2	8.28	4.90	72	8.95	6.03	95	-0.77	-.060
ITEM 3	0.10	0.53	72	0.11	0.55	95	-0.09	-.010
ITEM 4	0.01	0.12	72	0.08	0.40	95	-1.43	-.111
ITEM 5	0.12	0.53	72	0.16	0.76	95	-0.31	-.024
ITEM 6	1.36	1.70	72	1.73	2.25	95	-1.15	-.089
ITEM 7	1.60	2.33	72	1.00	1.76	95	+1.89	+.145
ITEM 8	5.44	5.75	72	4.13	4.13	95	+1.72	+.133
ITEM 9	6.19	5.14	72	6.06	4.99	95	+0.17	+.014
ITEM 10	0.72	2.22	72	0.37	1.45	95	+1.24	+.096
ITEM 11	7.10	6.97	72	8.76	8.08	95	-1.39	-.108
ITEM 12	3.65	5.87	72	2.29	4.42	95	+1.72	+.133
ITEM 13	0.44	1.78	72	0.32	1.87	95	+0.45	+.035
ITEM 14	0.56	3.93	72	0.86	3.46	95	-0.52	-.040
ONE STUD	15.76	9.27	72	14.86	8.33	95	+0.66	+.051
SM GROUP	4.00	7.31	72	2.54	4.71	95	+1.57	+.121
LG GROUP	4.25	8.05	72	3.68	7.53	95	+0.47	+.036
WHOL CLAS	12.01	10.97	72	14.34	10.80	95	-1.37	-.106

TABLE III-72

TEACHER CLASSROOM OBSERVATION
E/C OAKLAND, TIME 2

ITEMS & CONTEXT	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\bar{X}$	S.D.	N	$\bar{X}$	S.D.	N	t-val.	r_{pb}
ITEM 1	1.26	3.78	68	2.24	6.31	100	-1.14	-.088
ITEM 2	9.37	5.98	68	9.73	7.28	100	-0.34	-.026
ITEM 3	0.75	2.05	68	0.74	3.15	100	+0.02	+.000
ITEM 4	0.10	0.39	68	0.17	0.79	100	-0.65	-.050
ITEM 5	0.54	1.71	68	0.36	1.66	100	+0.70	+.054
ITEM 6	1.01	1.41	68	1.30	2.23	100	-0.94	-.073
ITEM 7	1.57	2.05	68	1.81	2.90	100	-0.58	-.045
ITEM 8	4.51	4.93	68	4.78	4.48	100	-0.36	-.028
ITEM 9	6.03	5.40	68	6.10	5.31	100	-0.08	-.000
ITEM 10	0.94	2.01	68	0.70	2.38	100	+0.69	+.053
ITEM 11	4.54	5.54	68	3.53	5.33	100	+1.19	+.092
ITEM 12	2.93	5.36	68	3.50	6.79	100	-0.58	-.045
ITEM 13	0.01	0.12	68	0.22	1.47	100	-1.14	-.088
ITEM 14	0.63	3.54	68	0.38	2.40	100	+0.55	+.042
ONE STUD	16.38	9.41	68	12.16	10.35	100	+2.69	+.204
SM GROUP	1.37	4.20	68	2.19	5.58	100	-1.03	-.080
LG GROUP	3.21	6.92	68	3.19	7.51	100	+0.01	+.000
WHOL CLAS	13.19	10.27	68	18.17	13.21	100	-2.62	-.199

instruction in the EXP school relative to the CON school did increase as predicted, the t-value going from +0.66 to +2.70. There was also a decline in "Whole Class" instruction in the EXP school, as predicted, although the trend is very slight, with the t-value going from -1.37 to -2.62.

The Teacher Classroom Observation results for Oakland are a mixture of consistent and erratic patterns. The consistent patterns are first that a majority of the changes were in the direction not predicted, and second, that this pattern is similar to what was observed of student classroom activities. However, there appears to be little congruence between student and teacher activities. For example, Oakland EXP students spent relatively more time "Answering Questions", but there is no corresponding increase in EXP teachers "Asking Questions". Furthermore, the student increase in "Listening" is not matched by a teacher increase in "Lecturing". On the other hand, there are some similarities. For example, the relative increase of EXP students in "Nonproductive" behavior is matched by increases in EXP teacher "Punishing" and "Administering". Thus, although negative results predominate in both the student and teacher classroom data in Oakland, there is no consistency between them.

In addition, the underlying causes of the changes vary in Oakland. In only two instances, "Criticizing" and "Small Group" instruction, were the t-value changes due to changes in the frequency of the activity in one school only. In the eight remaining areas where there was a discernible trend, it was the product of the combination of changes in both the EXP and CON schools. When this combining of changes is noted, the impact of incentives can be estimated in terms of what they did to trends in the EXP school as compared to trends at the CON school. Thus, in the CON school the context trend during the year was toward "Whole Class" and away from "One Student". In the EXP school, the tendency toward "Whole Class" was attenuated while the tendency away from "One Student" was reversed. Reversals of the tendencies established by the CON school are also found in "Listening" and "Supervising", both of which decreased in the EXP school; and in "Enriching", which increased in the EXP school. The

trends in these areas, therefore, may well be due to regression effects. Both the EXP and CON teacher activity in "Administrating" declined from Time 1 to Time 2, but the amount of decline was greater inthe CON school, which magnified the apparent impact in the EXP school.

These vicissitudes in the data inhibit any attempt to paint a detailed picture of the Oakland EXP school. Except for the substantial relative difference in the "One Student" context, which is not supported by the student observations, there is no safe generalization to be made about the impact of incentives on teacher activity. In support of the feeling that other sources of variation are at work (*e.g.*, the principal, historical factors, different cohorts), it is interesting to note that both the EXP and CON schools increased their frequency of "Testing" threefold from Time 1 to Time 2. This absolute and relative increase in the EXP school testing was expected and found in Cincinnati and Jacksonville. That it was found in the Oakland CON school is contrary to what was expected. However, the increase is not too surprising when one remembers that the State of California statewide achievement tests were to be administered soon after the observations were completed in April. This example is perhaps a bit too tidy and appropriate, but it does illustrate the problems of not understanding or accounting for many sources of variation in an evaluation of a project of this nature.

Tables III-73 and III-74 present the San Antonio Teacher Classroom Observation results. Among the 14 teacher activities, trends worth noting are found in five items. The trends were not in the direction predicted in two of these activities. Among these two reversals, one of the trends was rather substantial; in Item 3, "Lecturing", for which the t-value went from -1.84 to +1.29. A minor trend that was not predicted was found in Item 7, "Criticizing", where the t-value went from -1.73 to +0.05.

There were three trends in the direction hypothesized among the teacher activities with a substantial relative change in only one activity. In Item 8, "Listening", the difference in t-values from Time 1 to Time 2 was +2.71. Less substantial trends were found in Item 11, "Administrating",

TABLE III-73

TEACHER CLASSROOM OBSERVATION
E/C, SAN ANTONIO, TIME 1

ITEMS & CONTEXT	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
ITEM 1	1.10	4.98	60	0.04	0.28	52	+1.53	+.145
ITEM 2	5.87	6.68	60	5.38	5.60	52	+0.41	+.039
ITEM 3	1.07	3.23	60	2.31	3.90	52	-1.84	-.173
ITEM 4	0.00	0.00	60	0.02	0.14	52	UNDEF	UNDEF
ITEM 5	0.50	1.33	60	0.40	1.45	52	+0.37	+.035
ITEM 6	1.15	1.88	60	1.52	2.87	52	-0.81	-.077
ITEM 7	1.43	2.27	60	2.48	4.00	52	-1.73	-.163
ITEM 8	3.37	4.86	60	3.75	4.36	52	-0.44	-.041
ITEM 9	7.40	6.46	60	7.67	6.28	52	-0.23	-.022
ITEM 10	3.62	7.62	60	1.69	3.74	52	+1.66	+.156
ITEM 11	4.18	6.80	60	3.46	5.55	52	+0.61	+.058
ITEM 12	4.63	7.81	60	4.90	8.79	52	-0.17	-.017
ITEM 13	0.48	2.16	60	1.06	3.81	52	-1.00	-.095
ITEM 14	1.22	6.25	60	0.79	3.51	52	+0.44	+.041
ONE STUD	11.32	10.01	60	10.62	8.29	52	+0.40	+.039
SM GROUP	3.62	7.76	60	1.67	4.61	52	+1.58	+.149
LG GROUP	4.38	0.50	60	2.25	7.02	52	+1.33	+.126
WHOL CLAS	17.10	13.32	60	21.48	10.83	52	-1.89	-.177

TABLE III-74

TEACHER CLASSROOM OBSERVATION
E/C, SAN ANTONIO, TIME 2

ITEMS & CONTEXT	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
ITEM 1	0.00	0.00	72	0.43	2.53	60	UNDEF	UNDEF
ITEM 2	8.75	6.82	72	8.63	7.89	60	+0.09	+.010
ITEM 3	0.99	2.27	72	0.53	1.63	60	+1.29	+.112
ITEM 4	0.07	0.42	72	0.00	0.00	60	UNDEF	UNDEF
ITEM 5	0.00	0.00	72	0.05	0.39	60	UNDEF	UNDEF
ITEM 6	0.50	1.14	72	0.50	1.21	60	+0.00	+.000
ITEM 7	0.46	0.99	72	0.45	1.08	60	+0.05	+.000
ITEM 8	6.75	6.10	72	4.47	5.29	60	+2.27	+.195
ITEM 9	5.29	4.77	72	5.57	5.26	60	-0.31	-.028
ITEM 10	2.65	6.60	72	0.67	2.02	60	+2.24	+.193
ITEM 11	4.14	7.24	72	5.42	9.12	60	-0.90	-.079
ITEM 12	5.07	8.92	72	6.38	9.84	60	-0.80	-.070
ITEM 13	0.19	1.11	72	0.15	0.66	60	+0.27	+.024
ITEM 14	0.33	1.78	72	0.03	0.26	60	+1.29	+.112
ONE STUD	7.85	8.14	72	9.03	10.06	60	-0.75	-.066
SM GROUP	2.86	6.46	72	1.60	4.65	60	+1.26	+.110
LG GROUP	3.56	8.09	72	1.88	6.19	60	+1.32	+.114
WHOL CLAS	21.92	12.22	72	20.52	13.97	60	+0.61	+.054

with a t-value difference of -1.51; and "Peer Tutoring", with a t-value difference of +1.27.

There were two trends in the context of instruction in San Antonio, both in the direction not predicted. A substantial relative change was found in the "Whole Class" context, in which the status of the EXP school relative to the CON school changed from a Time 1 t-value of -1.89 to a Time 2 t-value of +0.61. This net difference of +2.50 means that the frequency of instruction to the total classroom increased in the EXP school relative to the CON school. The second trend, although smaller, was also not predicted. The frequency of "One Student" instruction decreases in the EXP school relative to the CON school, with a net t-value difference of -1.15.

A cursory examination of the means for Time 1 and Time 2 for both schools indicates that the sources of relative change appear to vary randomly. That is, in some cases the source of the trend revealed by the t-values is a large change in the CON school; in other cases it is a large change in the EXP school; and in others it arises out of the combination of slight changes in both schools. This has been found in the analysis of the other sites and is no surprise. However, in San Antonio the analysis is further complicated by the total absence of any recorded tallies for: the EXP school, Time 1, "Rewarding", and Time 2, "Punishing" and "Testing"; and for the CON school, Time 2, "Rewarding". The results presented here agree with the raw data submitted by the field representatives, but because there is no way of checking the accuracy of their work, these data are highly suspect.

Trends from Time 1 to Time 2 established by the CON school were reversed in the EXP school in two instances, and in both of these cases, "Testing" and "Rewarding", the missing data are a sufficient reason to dismiss them. In the remaining four areas where the trends are substantial, they are due to differences in rates of change between the two schools. "Punishing" and "One Student" instruction decline in both schools, but the decline is greater in the EXP school. "Listening" increases in both schools, but more so in the EXP school. The trend in "Administrating" is

due to an increase in this activity in the CON school, while it remains relatively stable in the EXP school. "Whole Class" instruction, on the other hand, increases in the EXP school but is constant from Time 1 to Time 2 in the CON school.

In San Antonio, the only impact of incentives revealed by this data is the trend toward "Whole Class" instruction. This is confirmed by the Student Observation results and is also supported by the relative tendency away from "One Student" instruction. Reversals in trends found in the CON school, which in any event could be a regression artifact, cannot be trusted because of the quality of the data in those items. For the other trends found, the EXP school differs not in kind but only very slightly in degree from the CON school.

Model-Based Comparisons--Teacher Classroom Observation Results

The results for the model-based comparisons are presented in Tables III-75 through III-82.

Tables III-75 and III-76 present the Time 1 and Time 2 results for the Teacher-Only sites, Cincinnati and Jacksonville. There would appear to be substantial trends in the direction predicted in five areas: Item 1, "Testing"; Item 2, "Explaining"; Item 6, "Praising"; Item 8, "Listening"; and Item 9, "Questionning". There would appear to be a change opposite to the direction hypothesized in "Large Group" instruction.

However, to conclude that these are the impacts of incentives on teacher classroom behavior would be to ignore the fact that the Jacksonville impacts predominate in the results. Only for Items 1, 2, and 6 are the trends similar for the two Teacher-Only sites. The trends in Items 8, 9, and "Large Group" are due to trends in Jacksonville only. Generalizations based upon an N of two are usually questionable with this type of exploratory data; but to generalize in this case about "Listening", "Questioning", and "Large Group" instruction, which are primarily attributable to Jacksonville and which are highly ambiguous in that context, would be misleading.

TABLE III-75

TEACHER CLASSROOM OBSERVATION
E/C, TEACHER ONLY, TIME 1

ITEMS & CONTEXT	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\bar{X}$	S.D.	N	$\bar{X}$	S.D.	N	t-val.	r_{pb}
ITEM 1	0.59	3.04	168	1.18	4.36	145	-1.40	-.079
ITEM 2	8.23	5.45	168	9.24	5.83	145	-1.59	-.090
ITEM 3	1.14	3.67	168	1.14	3.22	145	+0.01	+.000
ITEM 4	0.04	0.22	168	0.02	0.14	145	+0.71	+.040
ITEM 5	0.07	0.31	168	0.08	0.68	145	-0.18	-.010
ITEM 6	1.51	2.43	168	1.16	1.91	145	+1.42	+.080
ITEM 7	2.61	3.13	168	1.96	2.77	145	+1.93	+.109
ITEM 8	5.16	4.83	168	5.39	4.16	145	-0.45	-.026
ITEM 9	6.48	4.66	168	7.43	5.15	145	-1.70	-.096
ITEM 10	2.43	4.18	168	1.63	3.38	145	+1.83	+.103
ITEM 11	4.21	6.43	168	3.28	5.38	145	+1.37	+.077
ITEM 12	2.74	4.85	168	2.39	5.08	145	+0.63	+.036
ITEM 13	0.12	1.03	168	0.12	0.97	145	-0.04	-.000
ITEM 14	0.51	2.79	168	0.23	1.94	145	+1.03	+.058
ONE STUD	9.86	9.60	168	8.07	8.50	145	+1.74	+.098
SM GROUP	3.06	6.83	168	3.16	7.98	145	-0.12	-.000
LG GROUP	6.22	10.87	168	8.57	11.87	145	-1.83	-.103
WHOL CLAS	16.55	12.91	168	15.46	14.15	145	+0.71	+.040

TABLE III-76

TEACHER CLASSROOM OBSERVATION
E/C, TEACHER ONLY, TIME 2

ITEMS & CONTEXT	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\bar{X}$	S.D.	N	$\bar{X}$	S.D.	N	t-val.	r_{pb}
ITEM 1	0.81	3.92	182	0.31	1.72	180	+1.56	+.082
ITEM 2	7.68	5.45	184	6.53	5.83	180	+1.94	+.101
ITEM 3	1.71	3.42	184	1.31	3.90	180	+1.04	+.055
ITEM 4	0.02	0.18	184	0.01	0.15	180	+0.61	+.032
ITEM 5	0.05	0.29	184	0.08	0.43	180	-0.61	-.032
ITEM 6	1.35	1.95	184	0.58	1.37	180	+4.36	+.223
ITEM 7	1.95	2.82	184	1.52	2.69	180	+1.46	+.077
ITEM 8	5.14	5.18	184	4.13	4.72	180	+1.94	+.101
ITEM 9	6.66	5.28	184	6.20	6.23	180	+0.76	+.040
ITEM 10	1.56	3.12	184	1.27	2.79	180	+0.93	+.049
ITEM 11	2.40	5.09	184	2.18	4.32	180	+0.43	+.022
ITEM 12	2.78	4.71	184	2.19	4.78	180	+1.18	+.062
ITEM 13	0.11	0.69	184	0.27	1.69	180	-1.13	-.059
ITEM 14	0.08	0.63	184	0.17	1.08	180	-0.92	-.048
ONE STUD	8.88	9.21	184	6.04	8.01	180	+3.13	+.162
SM GROUP	3.03	7.47	184	2.79	6.94	180	+0.31	+.017
LG GROUP	7.59	11.77	184	5.34	10.12	180	+1.95	+.102
WHOL CLAS	12.79	13.11	184	12.53	13.60	180	+0.19	+.010

Tables III-77 and III-78 present the Teacher Classroom Observation results for the Parent-Teacher sites, Oakland and San Antonio. There are four trends worthy of any attention in these results, and they are all negative. In Item 1, "Testing"; Item 3, Lecturing", Item 12, "Supervising"; and the "Small Group" context, the trend is in the direction opposite to what was predicted. No conclusions can be drawn from this pattern because each trend is the result of the data from one site only. Thus, the apparent PT model trends toward less "Testing" and more "Lecturing" are not found in Oakland and are almost entirely the result of the San Antonio data. Similarly, the apparent trends toward "Supervising" and away from "Small Group" instruction can be traced only to Oakland. In fact, the results at the two sites, when considered and compared individually, are even more confusing. Thus, for the combined PT model data, one must conclude that there is no pattern in the results.

The next comparison to be considered is between the EXP schools in the Parent-Teacher model and the EXP schools in the Teacher-Only model. Tables III-79 and III-80 present the data for these schools. Although this comparison introduces at least regional and local uncontrolled sources of variance to complicate further attempts to discern patterns, it does provide a rough indication of the marginal differences between the two models. There are eight changes in the t-values from Time 1 to Time 2 that are greater than 2.00. Six of these changes are in the direction opposite to what was predicted: Item 6, "Praising", decreases in the PT model EXP schools relative to the TO model EXP schools; Item 9, "Questioning", decreases; Item 13, "Peer Tutoring", decreases; "Small Group" instruction decreases; and "Whole Class" instruction increases. In support of the hypotheses, one finds Item 2, "Explaining" and Item 4, "Rewarding" both increasing in the PT model EXP schools relative to the TO model EXP schools; while "Large Group" instruction decreases.

Based upon an examination of the means, the results in "Whole Class" and "Large Group" instruction, and Item 2, "Explaining" can probably be attributed more to regression effect than to any real impact because the models' means change in opposite directions. In the other five areas,

(Text resumes on page III-195)

TABLE III-77

TEACHER CLASSROOM OBSERVATION
E/C, PARENT TEACHER ONLY, TIME 1

ITEMS & CONTEXT	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
ITEM 1	0.76	3.76	132	0.54	2.67	147	+0.55	+.033
ITEM 2	7.18	5.88	132	7.69	6.10	147	-0.70	-.042
ITEM 3	0.54	2.25	132	0.88	2.57	147	-1.19	-.071
ITEM 4	0.01	0.09	132	0.06	0.34	147	-1.78	-.106
ITEM 5	0.30	0.99	132	0.24	1.06	147	+0.41	+.024
ITEM 6	1.27	1.78	132	1.65	2.48	147	-1.49	-.089
ITEM 7	1.52	2.30	132	1.52	2.85	147	-0.00	-.000
ITEM 8	4.50	5.44	132	3.99	4.20	147	+0.88	+.053
ITEM 9	6.74	5.78	132	6.63	5.51	147	+0.16	+.010
ITEM 10	2.04	5.56	132	0.84	2.58	147	+2.35	+.140
ITEM 11	5.77	7.02	132	6.88	7.70	147	-1.26	-.075
ITEM 12	4.10	6.81	132	3.21	6.41	147	+1.12	+.067
ITEM 13	0.46	1.95	132	0.58	2.73	147	-0.40	-.024
ITEM 14	0.86	5.11	132	0.83	3.47	147	+0.05	+.000
ONE STUD	13.74	9.83	132	13.36	8.54	147	+0.35	+.020
SM GROUP	3.83	7.49	132	2.23	4.67	147	+2.16	+.128
LG GROUP	4.31	8.70	132	3.18	7.36	147	+1.18	+.071
WHOL CLAS	14.33	12.31	132	16.86	11.30	147	-1.79	-.107

TABLE III-78

TEACHER CLASSROOM OBSERVATION
E/C, PARENT TEACHER ONLY, TIME 2

ITEMS & CONTEXT	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
ITEM 1	0.61	2.70	140	1.56	5.28	160	-1.92	-.110
ITEM 2	9.05	6.41	140	9.32	7.51	160	-0.33	-.020
ITEM 3	0.87	2.17	140	0.66	2.68	160	+0.74	+.042
ITEM 4	0.09	0.41	140	0.11	0.63	160	-0.33	-.020
ITEM 5	0.26	1.22	140	0.24	1.34	160	+0.14	+.010
ITEM 6	0.75	1.30	140	1.00	1.95	160	-1.29	-.074
ITEM 7	1.00	1.68	140	1.30	2.47	160	-1.21	-.070
ITEM 8	5.66	5.65	140	4.66	4.79	160	+1.66	+.096
ITEM 9	5.65	5.08	140	5.90	5.28	160	-0.42	-.024
ITEM 10	1.82	5.00	140	0.69	2.24	160	+2.59	+.148
ITEM 11	4.34	6.45	140	4.24	7.03	160	+0.13	+.010
ITEM 12	4.03	7.46	140	4.58	8.16	160	-0.61	-.035
ITEM 13	0.11	0.80	140	0.19	1.23	160	-0.71	-.041
ITEM 14	0.48	2.78	140	0.25	1.91	160	+0.84	+.049
ONE STUD	11.99	9.74	140	10.99	10.32	160	+0.86	+.050
SM GROUP	2.14	5.51	140	1.97	5.25	160	+0.27	+.014
LG GROUP	3.39	7.50	140	2.70	7.05	160	+0.82	+.047
WHOL CLAS	17.68	12.10	140	19.05	13.50	160	-0.92	-.053

TABLE III-79

TEACHER CLASSROOM OBSERVATION
PARENT TEACHER-E/TEACHER ONLY-E, TIME 1

ITEMS & CONTEXT	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
ITEM 1	0.76	3.76	132	0.59	3.04	168	+0.43	+.024
ITEM 2	7.18	5.88	132	8.23	5.45	168	-1.59	-.092
ITEM 3	0.54	2.25	132	1.14	3.67	168	-1.66	-.096
ITEM 4	0.01	0.09	132	0.04	0.22	168	-1.41	-.081
ITEM 5	0.30	0.99	132	0.07	0.31	168	+2.83	+.162
ITEM 6	1.27	1.78	132	1.51	2.43	168	-0.98	-.057
ITEM 7	1.52	2.30	132	2.61	3.13	168	-3.34	-.190
ITEM 8	4.50	5.44	132	5.16	4.83	168	-1.11	-.064
ITEM 9	6.74	5.78	132	6.48	4.66	168	+0.43	+.024
ITEM 10	2.04	5.56	132	2.43	4.18	168	-0.69	-.040
ITEM 11	5.77	7.02	132	4.21	6.43	168	+2.01	+.116
ITEM 12	4.10	6.81	132	2.74	4.85	168	+2.02	+.116
ITEM 13	0.46	1.95	132	0.20	1.03	168	+1.96	+.113
ITEM 14	0.86	5.11	132	0.51	2.79	168	+0.74	+.044
ONE STUD	13.74	9.83	132	9.86	9.60	168	+3.44	+.195
SM GROUP	3.83	7.49	132	3.06	6.83	168	+0.92	+.054
LG GROUP	4.31	8.70	132	6.22	10.87	168	-1.65	-.095
WHOL CLAS	14.33	12.31	132	16.55	12.91	168	-1.51	-.087

TABLE III-80

TEACHER CLASSROOM OBSERVATION
PARENT TEACHER-E/TEACHER ONLY-E, TIME 2

ITEMS & CONTEXT	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
ITEM 1	0.61	2.70	140	0.81	3.92	184	-0.51	-.028
ITEM 2	9.05	6.41	140	7.68	5.45	184	+2.08	+.115
ITEM 3	0.87	2.17	140	1.71	3.42	184	-2.53	-.139
ITEM 4	0.09	0.41	140	0.02	0.18	184	+1.90	+.105
ITEM 5	0.26	1.22	140	0.05	0.29	184	+2.26	+.125
ITEM 6	0.75	1.30	140	1.35	1.95	184	-3.14	-.172
ITEM 7	1.00	1.68	140	1.95	2.82	184	-3.52	-.192
ITEM 8	5.66	5.65	140	5.14	5.18	184	+0.87	+.049
ITEM 9	5.65	5.08	140	6.66	5.28	184	-1.73	-.096
ITEM 10	1.82	5.00	140	1.56	3.12	184	+0.58	+.000
ITEM 11	4.34	6.45	140	2.40	5.09	184	+3.02	+.166
ITEM 12	4.03	7.46	140	2.78	4.71	184	+1.84	+.102
ITEM 13	0.11	0.80	140	0.11	0.69	184	-0.08	-.000
ITEM 14	0.48	2.78	140	0.08	0.63	184	+1.88	+.104
ONE STUD	11.99	9.74	140	8.88	9.21	184	+2.94	+.162
SM GROUP	2.14	5.51	140	3.03	7.47	184	-1.19	-.066
LG GROUP	3.39	7.50	140	7.59	11.77	184	-3.70	-.202
WHOL CLAS	17.68	12.10	140	12.79	13.11	184	+3.44	+.188

the trends appear to be the result of an intensification in the PT model of patterns found in the TO model. For example, the EXP schools in both models tend away from "Small Group" instruction, but the rate of change is greater in the PT model. Again, it would appear that the net impact on teacher classroom behavior of offering incentives to parents in addition to the teachers is quite slight and more a matter of degree than kind. This is intuitive by reason, given the relatively high amount of physical and psychic isolation of classrooms from parental and community influences.

The final model-based comparison to be considered is a comparison of all of the EXP schools with all of the CON schools. Although the *caveat* about probable sources of random variation obtained in this comparison as in the previous model-based analyses, the trends uncovered can be useful if they are regarded as being highly tentative.

Tables III-81 and III-82 present the data for the four EXP schools and the four CON schools. Trends are found in five areas, three of which were in the direction hypothesized. In Item 2, "Explaining"; Item 6, "Praising'; and Item 8, "Listening", the t-value differences were greater than 2.00 and as predicted. However, contrary to what was predicted, the EXP teachers increased "Lecturing" and "Large Group" instruction. Overall, therefore, there are very few discernible changes in teacher behavior in the EXP schools. Furthermore, it would appear that where there are changes they combine to form a mixed pattern of traditional and innovative activities.

Teacher Classroom Observation Results--Discussion

If the line between art and science has not been crossed earlier in this report, and there is ample evidence that it has, then an attempt to summarize the results of the Teacher Classroom Observations provides that opportunity. Table III-83 summarizes the status of the hypotheses for the four sites in this project. The purpose of this project was not to prove or disprove theories or assumptions about the impact of incentives;

TABLE III-81

TEACHER CLASSROOM OBSERVATION
E/C, ALL, TIME 1

ITEMS & CONTEXT	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
ITEM 1	0.66	3.37	300	0.86	3.62	292	-0.68	-.028
ITEM 2	7.77	5.66	300	8.46	6.01	292	-1.44	-.059
ITEM 3	0.88	3.14	300	1.01	2.91	292	-0.54	-.022
ITEM 4	0.02	0.17	300	0.04	0.26	292	-0.99	-.040
ITEM 5	0.17	0.71	300	0.16	0.89	292	+0.09	+.000
ITEM 6	1.40	2.17	300	1.41	2.22	292	-0.02	-.000
ITEM 7	2.13	2.84	300	1.74	2.81	292	+1.68	+.069
ITEM 8	4.87	5.11	300	4.69	4.23	292	+0.47	+.020
ITEM 9	6.60	5.18	300	7.03	5.34	292	-1.00	-.041
ITEM 10	2.26	4.83	300	1.23	3.03	292	+3.08	+.126
ITEM 11	4.90	6.73	300	5.10	6.88	292	-0.36	-.014
ITEM 12	3.34	5.83	300	2.80	5.79	292	+1.12	+.046
ITEM 13	0.27	1.51	300	0.35	2.06	292	-0.56	-.022
ITEM 14	0.66	3.97	300	0.53	2.83	292	+0.47	+.020
ONE STUD	11.57	9.88	300	10.73	8.90	292	+1.08	+.045
SM GROUP	3.40	7.13	300	2.69	6.53	292	+1.25	+.052
LG GROUP	5.38	10.01	300	5.86	10.21	292	-0.57	-.024
WHOL CLAS	15.57	12.68	300	16.17	12.79	292	-0.57	-.024

TABLE III-82

TEACHER CLASSROOM OBSERVATION
E/C, ALL, TIME 2

ITEMS & CONTEXT	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
ITEM 1	0.73	3.45	324	0.90	3.88	340	-0.61	-.024
ITEM 2	8.27	5.91	324	7.84	6.81	340	+0.86	+.033
ITEM 3	1.35	2.97	324	1.00	3.39	340	+1.38	+.054
ITEM 4	0.05	0.30	324	0.06	0.45	340	-0.22	-.010
ITEM 5	0.15	0.83	324	0.16	0.97	340	-0.15	-.000
ITEM 6	1.09	1.72	324	0.78	1.68	340	+2.37	+.092
ITEM 7	1.54	2.44	324	1.42	2.59	340	+0.61	+.024
ITEM 8	5.36	5.39	324	4.38	4.75	340	+2.50	+.097
ITEM 9	6.22	5.21	324	6.06	5.80	340	+0.38	+.014
ITEM 10	1.67	4.03	324	1.00	2.56	340	+2.59	+.100
ITEM 11	3.23	5.79	324	3.15	5.84	340	+0.19	+.010
ITEM 12	3.32	6.07	324	3.32	6.69	340	+0.01	+.000
ITEM 13	0.11	0.74	324	0.23	1.49	340	-1.32	-.051
ITEM 14	0.25	1.89	324	0.21	1.53	340	+0.35	+.014
ONE STUD	10.23	9.55	324	8.37	9.48	340	+2.51	+.097
SM GROUP	2.65	6.70	324	2.41	6.21	340	+0.48	+.017
LG GROUP	5.77	10.34	324	4.10	8.90	340	+2.24	+.087
WHOL CLAS	14.90	12.90	324	15.60	13.92	340	-0.67	-.026

TABLE III-83

TEACHER CLASSROOM OBSERVATION RESULTS
COMPARISON OF HYPOTHESES AND RESULTS

CONTEXT/ ITEM	ACTIVITY	HYP	RESULTS			
			CIN	JAX	OAK	SAN
1	Testing	E > C	+	+		
2	Explaining	E > C	+	+		
3	Lecturing	E < C		-		-
4	Rewarding	E > C				
5	Punishing	E < C	-		-	
6	Praising	E > C	+	+		
7	Criticizing	E < C	-	+	+	-
8	Listening	E > C		+	-	+
9	Questioning	E > C		+		
10	Drilling	E < C				
11	Administrating	E < C		+	-	+
12	Supervising	E > C	+		-	
13	Peer Tutoring	E > C			-	+
14	Enriching	E > C		-	+	
	One Student	E > C	+		+	-
	Small Group	E > C		+	-	
	Large Group	E < C		-		
	Whole Class	E < C			+	1

KEY:

+ = Results are in the direction hypothesized.
- = Results are not in the direction hypothesized.

consequently, this portion of the analysis was focused upon revealing trends within the variables and then discerning patterns among them. Thus, the symbols used in Table III-83, do not represent the results of formal tests of significance.

There were four opportunities to examine 18 hypotheses. Among the 72 cells, some trend was indicated in 35. Of those 35, 20 of the trends were in the direction predicted. Thus, trends in support of the hypotheses were found in slightly over half of the cases where any trend existed (20/35 = 60%). There are many probable sources of error in the instrument and procedures that have been discussed in the Student Classroom Observation results, and those factors are acknowledged in these results as well. However, the depressingly low percentages of both the findings and confirmations must also reflect on what is known, or thought to be known, in any organized way about teaching behavior.

Examining the variables across sites, one must conclude that there is very little evidence of consistent trends due to incentives. "Testing" increased in the two Teacher-Only sites, but an examination of the EXP and CON school means in both sites would indicate that regression effects are operative in both cases. In any event, no similar relative increase was found in Oakland. It would be hazardous, however, to contrast the two models in this activity and to generalize from this comparison. The San Antonio data are suspect; and in fact, there were very substantial increases in "Testing" in both schools in Oakland, attributable, it is assumed, to the California testing program.

The relative frequency of "Explaining" in Cincinnati and Jacksonville is the result of declines in this activity in both schools at both sites. However, the magnitude of the decline was less in both EXP schools. "Lecturing" increased in two sites, which was not expected; however, only in Jacksonville, where the frequency of the activity doubled, can this trend be vigorously defended. On the other hand, this trend is in apparent contradiction to the concomitant increases in "Explaining" and "Listening" in Jacksonville.

Items 4, 5, 6, and 7 are logically related in that "Rewarding" and "Praising" are the reverse of "Punishing" and "Criticizing". Contrasting these two pairs of items yields no firm conclusions except that the teachers appear to use different motivational styles about equally. Thus, no clear instructional ideology emerges, even in those three sites that evidenced a trend toward individualized instruction (Cincinnati, Jacksonville, and Oakland). Considered together because they imply a commitment to innovative teaching, the results of "Supervising", "Peer Tutoring", and "Enriching" would indicate a very slight tendency away from that style in three of the four sites.

Across the sites, therefore, the results are ambiguous as far as teacher activities are concerned. Furthermore, there does not appear to be any tendency to conduct instruction in smaller or larger groups. In both Cincinnati and Oakland, the apparent focus upon "One Student" may be due to regression. The same artifact may be found in both the "Small Group" and "Large Group" results in Jacksonville. As with their classroom activities, the teachers appear to have only a very slight preference for individual instruction. However, regression effects are so probable in this trend that it must be advanced very cautiously.

Within each of the sites or the two models, an examination of the results is equally unprofitable. In Cincinnati, the EXP teachers appear to combine "Punishing" and "Criticizing" along with "Praising". In Jacksonville, the consistency of predicted behavior over "Praising", "Criticizing", and "Listening" is contrary to the decline in "Enriching" and the increase in "Lecturing". The apparent consistency between the two sites in "Testing", "Explaining", and "Praising" are probably not due to the particular incentives model found there but rather are attributable to site anomalies. Thus, they can only be noted in this analysis.

In Oakland, the EXP teachers are relatively more frequently "Punishing", but they also decline in "Criticizing". It would appear also, that the peripherals, "Supervising" and "Peer Tutoring" decline somewhat. However, there is a counter trend in the relative increase in "Enriching" and "One Student" instruction. In San Antonio, the more traditional activities of

"Lecturing" and "Criticizing" increase, contrary to expectations; but so do "Listening" and "Peer Tutoring", a rather innovative practice. As an appropriate illustration of the dangers of drawing conclusions about the impact of the Parent-Teacher model upon the classroom, it should be noted that in those items or contexts where tendencies were found in the two sites the trends are always reversed. Thus, "Criticizing" increases in Oakland, but declines in San Antonio; "Peer Tutoring" declines in Oakland but increases in San Antonio; and so on. As with the Student Classroom Observation results, the absence of any consistent pattern of trends is a consistent pattern in these results.

Analysis of Parent Attitudes and Behavior

Introduction

The preceding sections of this chapter have discussed results obtained from the various instruments designed to measure the attitudes and behavior of students and teachers. The results to be reported in this section will deal with the third major target population, the parents. Since this project was one in which one of the treatments involved parents directly, there is a particular interest in the changes that might be observed in their attitudes or behavior as a result of their being made eligible for cash payments. However, even aside from the direct impact of the incentives model in the two sites where parents could receive payment, there is also the related question as to the role played by parents when an educational innovation is introduced to their school. Thus, a second reason for interest in the possible impact of the project on parents is to see whether there is a spillover effect on parent attitudes and behavior, even in the sites where only the teachers actually were eligible for cash payment.

To tap the attitudes and behaviors of the parents, two instruments were used. The first of these was a Parent Questionnaire, designed to be completed at two times during the year by all parents in the CON schools as well as the EXP schools in all four sites. Due to delays in obtaining forms clearance, these Questionnaires were not available for use until March. Thus, as with the Student Questionnaire and the Teacher Questionnaire, the first wave of Parent Questionnaires was administered in March. The Questionnaires were given to each child at school, and the child was asked to take them home, and to ask his parents to complete them and mail them back in an envelope which was provided, directly to the TAC office in Washington. The second-wave Questionnaire to the parents was administered in a similar fashion.

At the TAC office, the Parent Questionnaires were compiled as they came in, and were allowed to be included in the analysis until early

in July, at which time the Questionnaires which had been received were keypunched for the analysis processing. If it is assumed that the number of eligible students is 4042 (which is the average of the number of Student Questionnaires received on the two waves), then the response rate for the Parent Questionnaire is as follows. There were 1329 usable Parent Questionnaires received at the TAC office from the first wave, for a response percentage of 32.9%. There were 801 usable Parent Questionnaires received at the second wave, for a response percentage of 19.8%. These low response rates limit the clarity of interpretations that can be made of the results from the Parent Questionnaire data. It is reasonable to suspect that the parents who were more favorably inclined toward the project would more frequently take the trouble to respond to the request that they fill out a questionnaire. Also, it is equally reasonable to suspect that the parents who did return the Questionnaire are in general more educated and affluent and academically supportive than those who did not. Consequently, the results from this analysis must be interpreted with caution.

The second instrument used to collect information from the parents was a Parent Interview Schedule. Interviews were carried out with approximately 200 parents throughout the project at two times during the year. The first of these interviews was conducted in December, and 206 interviews were completed. The second interview was conducted during May, and 189 were completed. The sampling procedure used for choosing the parents for these interviews was the same used for choosing the students. That is, a random sample of approximately 25 students was chosen to be interviewed in each of the eight schools; then their parents were interviewed; and then the students were interviewed.

The content of the two instruments used with the parents includes questions asking about attitudes and opinions, on the one hand, and also questions asking about the typical behavior of the parent, and, occasionally, the behavior of their child. Because these items actually are reports of behavior rather than objectively observed behavior, they have been kept together in building the various indexes, and will be treated together in the discussion.

Index Construction

As in the previous instruments, the content of the Parent instruments was determined by identifying various attitudes and behaviors which might be expected to be influenced by the introduction of the incentives models. Consequently, the indexes were constructed by combining all questions which dealt with a single broad topic, and treating the responses to each of these questions as components of a single index. There was a close but not perfect parallelism between the indexes developed from the Parent Interview and several of the indexes developed from the Parent Questionnaire. Thus, these indexes should be regarded as reflecting very similar, although not quite identical, content.

The Parent Questionnaire contained 38 numbered items. From this instrument, 12 indexes were constructed and used in the analysis. The names of these indexes, and the items composing each, are presented in Table III-84.

Similarly, there were six indexes constructed from the 20 items of the Parent Interview. These index names, and the items composing each index, are presented in Table III-85.

Single-Site Comparisons--Parent Questionnaire and Interview Results

The first site to be discussed is Cincinnati. The results for Cincinnati are presented in Table III-86. They reflect in general a somewhat negative change in the EXP school relative to the CON school between the first and second administrations of the instruments. From the Parent Questionnaire, only two indexes show a change large enough to note between Time 1 and Time 2. These are the index for "Adult Help for Student" and the index for "Parent-Teacher Contact". For both of these indexes, the direction of the change is negative; that is, contrary to what had been hypothesized. For the index of "Adult Help for Student", the change is due to a decline in the level of adult help at the EXP school in Cincinnati, and a very slight increase in the level of adult

(Text resumes on page III-210)

TABLE III-84

INDEXES CONSTRUCTED FROM THE PARENT QUESTIONNAIRE

1. Help for Student

Question No.	
19	Your child is helped in math by someone else at home
20	Your child is helped in math by an outside adult
25	Your child is helped in reading by someone else at home
26	Your child is helped in reading by an outside adult

2. Parent-Teacher Contact

Question No.	
34	Number of meetings with child's teacher in past two months
35	Number of phone conversations with child's teacher in past two months
36	Number of written notes from child's teacher in past two months

3. Parent-Parent Contact

Question No.	
37	Number of meetings of parents attended at school in past two months
38	Number of meetings of parents attended not at school in past two months

4. Parent Attitude toward Child's Achievement

Question No.	
17	How well is your child doing in math...(poorly)...(well)
18	How well is your child doing in reading...(poorly)...(well)

5. Parent Interest in Child's Achievement

Question No.	
9	To me, my child's ability to do math well...(does not matter)...(is very important)
10	To me, my child's ability to read well...(does not matter)...(is very important)

6. Parent Attitude toward School

Question No.	
1	My child's school...(could stand improvement)...(could serve as a model to others)
3	How do you feel about your child's school...(dissatisfied)...(satisfied)
7	My feeling about discipline at my child's school is...(dissatisfied)...(satisfied)

7. Parent Attitude toward Teacher

Question No.	
2	My child's teacher is...(one of the worst in the school)...(one of the best in the school)
6	When it comes to my child's education, his teachers...(do not care)...(do care alot)
11	How do you feel about your child's teacher...(dissatisfied)...(satisfied)

-Continued

TABLE III-84 (Cont'd.)

7. Parent Attitude toward Teacher (Cont'd.)

Question No.	
13	How is your child's teacher doing in teaching arithmetic to him ...(poor job)...(good job)
14	How is your child's teacher doing in teaching reading to him... (poor job)...(good job)

8. Students Help Students

Question No.	
22	How often does your child get help with arithmetic from another child...(never)...(about every day)
23	How often does your child give help with arithmetic to another child...(never)...(about every day)
28	How often does your child get help with reading from another child...(never)...(about every day)
29	How often does your child give help in reading to another child ...(never)...(about every day)

9. Parent Attitude about Motivation

Question No.	
5	A good way to get my child to do well in school is to punish him when he does poorly...(strongly agree)...(disagree)
15	A good way to get my child to do well in school is to praise him when he does well...(disagree)...(agree)

10. Adults Help Neighbor's Children

Question No.	
21	How often does an adult in your home help other people's children with arithmetic...(never)...(about every day)
27	How often does an adult in your home help other people's children with reading...(never)...(about every day)

11. Parent Attitude toward Incentives to Teachers

Question No.	
4	Offering teachers more money if their pupils learn more is... (not likely to work)...(likely to work)
12	Offering teachers more money if their pupils learn more is... (not proper)...(proper)

12. Parent Attitude toward Incentives to Parents

Question No.	
8	Encouraging parents with money if their child's class learns more is...(not proper)...(proper)
16	Offering parents money if their child's class learns more is ...(not likely to work)...(likely to work)

TABLE III-85

INDEXES CONSTRUCTED FROM THE PARENT INTERVIEW

1. Adult Help for Student

Question No.	
4	Do you help your child in doing his/her homework
9	Do you get extra reading and/or arithmetic learning materials for your child
15	Do any adults outside of your home help your child with his/her homework

2. Parent-School Contact

Question No.	
2	Do you attend school activities, e.g, PTA meetings
12	Do you think parents need to be familiar with their child's teacher
13	Have you visited or met with your child's teacher
20	Have you discussed or attended meetings about the payments of money project with your child's teacher or other school official

3. Parent-Parent Contact

Question No.	
5	Do you think that groups of parents working together can improve their children's school grades
6	Have you organized any extra help sections for reading/arithmetic work for your children, with other parents in the neighborhood
19	Have you discussed or been to meetings with other parents about the payments of money project

4. Parent Satisfaction with Child's Achievement

Question No.	
7	Are you satisfied with how your child is doing in math
8	Are you satisfied with how your child is doing in reading

5. Parent Supports Child's Learning

Question No.	
3	How often do you encourage your child to do well at his school work
4	Do you help your child in doing his/her homework
9	Do you get extra reading and/or arithmetic learning material for your child
10	Do you encourage your child to use the public library

6. Parent Report of Student Learning Habits

Question No.	
11	How often does your child use the public/school library
14	Do your child's friends work with your child on their reading/arithmetic homework

TABLE III-86-1

PARENT QUESTIONNAIRE AND INTERVIEW RESULTS
E/C, CINCINNATI, TIME 1

INDEX	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\bar{X}$	S.D.	N	$\bar{X}$	S.D.	N	t-val.	r_{pb}
PQ Help from Adults	12.98	3.72	121	12.30	4.02	91	+1.27	+.087
PQ Parent-School Contacts	2.43	4.86	162	2.49	6.27	132	-0.10	-.000
PQ Parent-Parent Contacts	0.67	1.50	163	0.78	1.50	133	-0.64	-.037
PQ Satis w/Child's Ach	5.11	1.32	166	5.14	1.24	131	-0.20	-.010
PQ Care about Child's Ach	5.98	0.22	165	5.95	0.37	133	+0.61	+.035
PQ Satis w/School	6.03	1.97	162	6.56	1.79	125	-2.34	-.137
PQ Satis w/Teacher	12.81	2.53	159	13.58	1.85	128	-2.87	-.168
PQ Peer Tutoring	11.56	5.09	78	11.54	5.87	65	+0.03	+.000
PQ Discipline	4.98	1.10	165	4.95	1.03	135	+0.22	+.014
PQ Help Neighbor's Child	3.43	2.08	128	3.55	2.35	104	-0.41	-.026
PQ Attit Re Paying Teachers	3.59	1.49	163	3.91	1.53	132	-1.81	-.105
PQ Attit Re Paying Parents	3.07	1.30	166	2.99	1.21	134	+0.50	+.028
PI Help from Adults	5.72	2.75	25	6.00	2.67	22	-0.35	-.053
PI Parent-School Contact	9.24	4.24	25	7.50	2.48	22	+1.69	+2.44
PI Parent-Parent Contact	4.88	2.01	25	5.59	2.20	22	-1.16	-.170
PI Satis w/Child's Ach	7.44	2.72	25	6.77	3.01	22	+0.80	+.118
PI Encourage Child	9.96	3.35	25	10.27	3.24	22	-0.32	-.048
PI Student Learning Behav	3.44	1.76	25	4.09	1.85	22	-1.24	-.181

TABLE III-86-2

PARENT QUESTIONNAIRE AND INTERVIEW RESULTS
E/C, CINCINNATI, TIME 2

INDEX	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\bar{X}$	S.D.	N	$\bar{X}$	S.D.	N	t-val.	r_{pb}
PQ Help from Adults	11.82	3.87	44	12.53	4.66	51	-0.80	-.083
PQ Parent-School Contacts	1.83	2.75	70	4.58	10.00	59	-2.20	-.192
PQ Parent-Parent Contacts	0.90	2.57	69	1.43	3.52	60	-0.99	-.088
PQ Satis w/Child's Ach	5.09	1.46	70	5.29	1.30	65	-0.87	-.075
PQ Care about Child's Ach	5.97	0.17	71	6.00	0.00	65	UNDEF	UNDEF
PQ Satis w/School	5.86	2.06	64	6.72	1.92	64	-2.44	-.212
PQ Satis w/Teacher	13.24	2.11	67	14.07	1.65	59	-2.43	-.213
PQ Peer Tutoring	11.09	4.93	35	10.00	6.12	33	+0.81	+.099
PQ Discipline	4.81	1.01	70	4.70	1.18	64	+0.59	+.051
PQ Help Neighbor's Child	4.03	2.39	46	3.77	2.58	52	+0.50	+.051
PQ Attit Re Paying Teachers	3.64	1.44	70	4.02	1.56	64	-1.44	-.124
PQ Attit Re Paying Parents	3.04	1.17	70	2.95	1.03	64	+0.45	+.041
PI Help from Adults	6.20	2.89	20	5.73	2.99	22	+0.52	+.082
PI Parent-School Contact	7.10	3.35	20	8.55	3.86	22	-1.29	-.200
PI Parent-Parent Contact	4.50	2.52	20	4.14	1.88	22	+0.53	+.084
PI Satis w/Child's Ach	7.25	3.06	20	7.68	2.19	22	-0.53	-.084
PI Encourage Child	9.55	4.11	20	9.91	4.25	22	-0.28	-.044
PI Student Learning Behav	3.35	2.18	20	4.00	1.95	22	-1.02	-.159

help at the CON school in Cincinnati. On the other index, that for "Parent-Teacher Contact", the observed net shift is the result of a combination of a decline at the EXP school, and a larger rise of such contact at the CON school. Thus, the EXP school begins with an average score on this index at Time 1 of 2.43, and has declined to a score of 1.83 at the Time 2 administration. However, the CON school, which began at approximately the same level, 2.49, rises to an average of 4.58 at the Time 2 administration. This result seems once again to confirm that in Cincinnati, the CON school was making real efforts to improve and operate at the very best level possible.

The indexes constructed from the Parent Interview for Cincinnati confirm the shift toward relatively less contact between the parents and the school staff in the EXP school. The definitions of these two indexes are not quite identical, and the index from the interview emphasizes contact not just with the teacher, but with other aspects of the school as well. Nonetheless, the two indexes should be closely comparable. An inspection of Table III-86 indicates that indeed there is a notable net shift on the Interview-based index, "Parent-School Contact", in the direction opposite to the hypothesis (but consistent with the corresponding result from the Questionnaire-based index). Further inspection of the means in the table indicates that this pattern is attributable to the mixture of a large drop in contact rate at the EXP school and a smaller rise in contact rates at the CON school. Thus, both instruments suggest the same pattern.

The Interview-based index for "Parent-Parent Contact" at the EXP school shows an increase relative to that of the CON school. The t-value at Time 1 for this index is -1.16, and at Time 2 the corresponding t-value is +0.53. This change occurs through a type of lesser deterioration process. That is, the level of parent contact with other parents, as reflected in the Interview-based index, actually declines between Time 1 and Time 2 from an average of 4.88 to an average of 4.50. However, the CON school suffers a greater decline, from a Time 1 average of 5.59 to a Time 2 average of 4.14. In other words, this apparently positive result in the EXP school is really the lesser of two evils.

A third index based on the Parent Interview which shows a notable change from Time 1 to Time 2 is that for the "Parent Satisfaction with Child's Achievement". This index, too, is not quite identical to the one on the Parent Questionnaire. The Interview-based index refers specifically to "satisfaction", and the index based on the Questionnaire is phrased in terms of general perception of how well the child is doing in his school work. The t-value for this index based on the Interview changes from +0.80 at Time 1, indicating that the parents in the EXP school are more satisfied, to a value of -0.53, indicating that at Time 2 the parents in the CON school are more satisfied. This change comes about, however, primarily through a rise from 6.77 to 7.68 in the average for the CON school. The EXP school is nearly constant at the two times; its Time 1 average being 7.44 and its Time 2 average being 7.25.

The changes observed in the t-values in Cincinnati are not of the sort that are likely to be due to regression toward the mean. Thus, it seems reasonable to draw the tentative conclusion that the impact of the incentives model on parent attitudes and behaviors in Cincinnati was not positive, and possibly was negative. Of course, in interpreting these results as indicating a negative impact of the incentive treatment, one is ignoring the peculiar history of the project in Cincinnati. At that site, there was a prolonged absence of the EXP school principal, and an expression of an unusually high motivation at the CON school.

The second of the individual sites is Jacksonville. The results of the Parent Questionnaire and Parent Interview at Jacksonville are presented in Table III-87. They indicate a generally negative, although somewhat mixed, pattern. Of the 12 indexes constructed from the Parent Questionnaire, there are two which show a notable shift in the predicted direction, and five which show a notable shift in the direction contrary to that predicted by the hypothesis. Similarly, of the six indexes constructed from the Parent Interview, there is only one which shows a positive shift, and two which show a negative shift in the Jacksonville data.

TABLE III-87-1

PARENT QUESTIONNAIRE AND INTERVIEW RESULTS
E/C, JACKSONVILLE, TIME 1

INDEX	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
PQ Help from Adults	13.51	4.49	140	13.03	4.02	74	+0.78	+.054
PQ Parent-School Contacts	2.70	7.68	189	1.89	2.33	104	+1.05	+.062
PQ Parent-Parent Contacts	1.01	2.56	193	0.92	1.28	104	+0.33	+.020
PQ Satis w/Child's Ach	5.29	1.21	198	4.84	1.51	105	+2.70	+.154
PQ Care about Child's Ach	5.99	0.07	197	5.98	0.14	106	+1.16	+.066
PQ Satis w/School	6.60	1.58	191	6.16	1.57	101	+2.29	+.133
PQ Satis w/Teacher	13.59	2.22	184	12.87	2.34	101	+2.57	+.151
PQ Peer Tutoring	12.51	5.57	90	10.84	5.23	45	+1.67	+.144
PQ Discipline	4.78	1.13	196	4.67	1.10	106	+0.82	+.047
PQ Help Neighbor's Child	4.28	2.62	135	4.19	2.69	75	+0.25	+.017
PQ Attit Re Paying Teachers	3.95	1.46	197	3.42	1.39	104	+3.02	+.172
PQ Attit Re Paying Parents	3.11	1.25	196	2.66	1.02	106	+3.15	+.179
PI Help from Adults	7.78	2.04	27	8.76	2.59	21	-1.47	-.212
PI Parent-School Contact	7.52	3.74	27	8.43	3.99	21	-0.81	-.119
PI Parent-Parent Contact	6.41	1.08	27	5.05	2.01	21	+3.00	+.405
PI Satis w/Child's Ach	7.04	2.85	27	8.33	2.38	21	-1.67	-.240
PI Encourage Child	12.22	2.39	27	12.90	2.57	21	-0.95	-.139
PI Student Learning Behav	4.93	1.90	27	5.14	1.11	21	-0.46	-.069

TABLE III-87-2

PARENT QUESTIONNAIRE AND INTERVIEW RESULTS
E/C, JACKSONVILLE, TIME 2

INDEX	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
PQ Help from Adults	13.33	4.38	85	13.20	3.72	60	+0.19	+.014
PQ Parent-School Contacts	2.62	4.08	113	2.06	4.18	90	+0.97	+.068
PQ Parent-Parent Contacts	1.00	1.49	112	0.99	1.36	92	+0.05	+.000
PQ Satis w/Child's Ach	5.12	1.41	128	5.05	1.40	100	+0.40	+.026
PQ Care about Child's Ach	5.97	0.22	127	5.98	0.21	94	-0.35	-.024
PQ Satis w/School	7.34	1.59	117	5.99	1.81	93	+5.76	+3.71
PQ Satis w/Teacher	13.67	1.91	120	13.40	2.17	89	+0.93	+.064
PQ Peer Tutoring	11.86	5.70	70	12.12	5.55	48	-0.25	-.024
PQ Discipline	4.93	1.09	125	5.06	0.97	95	-0.95	-.065
PQ Help Neighbor's Child	4.51	2.76	92	3.78	2.26	64	+1.75	+.139
PQ Attit Re Paying Teachers	4.11	1.39	125	3.66	1.45	92	+2.31	+.156
PQ Attit Re Paying Parents	3.47	1.44	126	2.76	1.13	95	+3.98	+.259
PI Help from Adults	7.46	1.90	28	8.43	2.20	21	-1.64	-.233
PI Parent-School Contact	8.29	3.26	28	9.43	6.82	21	-1.13	-.162
PI Parent-Parent Contact	3.46	1.35	28	3.24	0.89	21	+0.67	+.097
PI Satis w/Child's Ach	8.04	3.00	28	7.76	2.53	21	+0.34	+.049
PI Encourage Child	12.36	1.75	28	13.14	1.88	21	-1.51	-.215
PI Student Learning Behav	4.68	1.42	28	5.43	1.33	21	-1.89	-.265

Of the Questionnaire-based indexes, the negative shifts occur on the following: "Parent Attitude toward Child's Achievement", "Parent Interest in Child's Achievement", "Parent Attitude toward Teacher", "Students Help Students", and "Parent Attitude about Motivation". The observed shift in the t-values for "Parent Attitude toward Child's Achievement", from a Time 1 t-value of +2.70 to a Time 2 t-value of +0.40, is caused by a rise in the average at the CON school from 4.84 to 5.05, and a small decline at the EXP school from 5.29 to 5.13. The shift is one that could be due to a regression toward the mean. Similarly, the index "Parent Interest in Child's Achievement" shows a change from a Time 1 t-value of +1.16 to a Time 2 t-value of -0.35. This change arises from the small variance of the index, which in all four cases has averages greater than 5.96, with a logical ceiling of 6.00. Thus, this result must be regarded as due probably to chance fluctuation and the very small observed variance on these items.

Similarly, the shifts observed on the other indexes are due to complex causes. For instance, the shift on the index "Parent Attitude toward Teacher" from a t-value of +2.57 to a t-value of +0.93 is due almost entirely to a rise in the attitude held by parents of children in the CON school, from an average of 12.87 at Time 1, to an average of 13.40 at Time 2. The change in the "Students Help Students" index is due to shifts in both schools. The EXP school shifts down and the CON school shifts up, thus making the net pattern on which the t-value goes from +1.67 at Time 1 to -0.25 at Time 2. The last of the Questionnaire-based indexes which show substantial shift in t-value is that for "Parent Attitude about Motivation". On this index, parents in both schools are increasing in positive attitude over time, but those in the CON school show a greater increase.

Turning now to the indexes derived from the Parent Interview in Jacksonville, one finds that there are shifts contrary to the predicted direction on the indexes "Parent-Parent Contact" and "Parent Report of Student Learning Habits". The change on the "Parent-Parent Contact" index is one which might well be due to a regression effect. That is, the

t-value at Time 1 is -3.00, and at Time 2 is +0.67. On this index, both schools decline considerably from Time 1 to Time 2, but the EXP school declines more. The average at the EXP school on the "Parent-Parent Contact" index at Time 1 is 6.41 and at Time 2 the corresponding average has shrunk to 3.46. In the CON school, the Time 1 average is 5.05, and it shrinks to 3.24 at Time 2.

The observed shift in the "Parent Report of Student Learning Behavior" index occurs through two cumulative changes: an increase over time at the CON school and a decrease over time at the EXP school. Finally, the positive change that was observed on the Interview-based indexes occurs on the index for "Parent Satisfaction with Child's Achievement". On this index, the t-value changes from -1.67 at Time 1 to +0.34 at Time 2. Thus, this could be due to a regression effect.

Despite the varied sources of these differences, it does seem fair to conclude that the impact of the incentives model on parent attitudes and behavior in Jacksonville was more negative than positive.

The third individual site to be examined is Oakland. The results for Oakland are presented in Table III-88. These results indicate a generally positive impact. There are four differences in the predicted direction on the Questionnaire-based indexes, and two differences in the direction opposite to what was predicted. On the Interview-based indexes, there are four positive differences, and no difference contrary to prediction. Interestingly, the pattern of results indicates a general improvement of the parents' attitudes about their child and the school, but the two instances of a prediction contrary to hypothesis are both on the attitudes toward the payment of incentives.

As can be seen from Table III-88, the index for "Parent-Parent Contact" has a t-value at Time 1 of +0.31, and a t-value at Time 2 of +2.28. This increase is the result of a considerably greater increase in the rate of such contact at the EXP school, although there is also an increase at the CON school. The average on this index at the EXP school at Time 1 is 5.68, and at Time 2 is 8.12. The average at the CON school at Time 1 is 5.54, and at Time 2 has risen only to 6.92.

TABLE III-88-1

PARENT QUESTIONNAIRE AND INTERVIEW RESULTS
E/C, OAKLAND, TIME 1

INDEX	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
PQ Help from Adults	13.82	3.95	108	13.44	4.09	119	+0.72	+.048
PQ Parent-School Contacts	3.08	4.42	157	2.08	3.73	154	+2.15	+.121
PQ Parent-Parent Contacts	1.35	2.07	158	1.08	2.04	153	+1.16	+.066
PQ Satis w/Child's Ach	4.89	1.44	160	5.38	1.15	153	-3.29	-.183
PQ Care about Child's Ach	5.96	0.36	164	6.00	0.00	151	UNDEF	UNDEF
PQ Satis w/School	5.81	1.90	155	5.99	1.89	139	-0.81	-.048
PQ Satis w/Teacher	13.03	2.44	154	13.56	1.90	146	-2.12	-.122
PQ Peer Tutoring	12.24	5.41	88	11.68	4.90	87	+0.72	+.055
PQ Discipline	4.97	1.05	162	5.11	1.03	151	-1.16	-.066
PQ Help Neighbor's Child	4.06	2.66	120	3.88	2.42	126	+0.55	+.035
PQ Attit Re Paying Teachers	3.99	1.49	160	3.36	1.41	148	+3.80	+.212
PQ Attit Re Paying Parents	3.56	1.46	160	2.68	0.95	149	+6.22	+.335
PI Help from Adults	8.18	3.30	38	8.62	3.02	26	-0.53	-.067
PI Parent-School Contact	10.63	4.06	38	10.35	2.98	26	+0.31	+.039
PI Parent-Parent Contact	5.68	2.03	38	5.54	1.65	26	+0.30	+.039
PI Satis w/Child's Ach	7.50	2.37	38	8.65	1.67	26	-2.14	-.263
PI Encourage Child	13.05	3.35	38	12.38	2.93	26	+0.82	+.104
PI Student Learning Behav	4.32	1.74	38	4.27	1.48	26	+0.11	+.014

TABLE III-88-2

PARENT QUESTIONNAIRE AND INTERVIEW RESULTS
E/C, OAKLAND, TIME 2

INDEX	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\bar{X}$	S.D.	N	$\bar{X}$	S.D.	N	t-val.	r_{pb}
PQ Help from Adults	13.52	3.81	79	13.04	4.49	56	+0.67	+.058
PQ Parent-School Contacts	3.35	5.30	102	2.01	5.01	72	+1.68	+.127
PQ Parent-Parent Contacts	1.22	2.01	101	0.57	1.15	74	+2.50	+.187
PQ Satis w/Child's Ach	5.29	1.26	96	5.38	1.18	77	-0.45	-.035
PQ Care about Child's Ach	5.99	0.10	103	6.00	0.00	76	UNDEF	UNDEF
PQ Satis w/School	5.99	1.97	93	6.16	1.95	77	-0.55	-.042
PQ Satis w/Teacher	13.66	2.00	93	13.68	1.97	74	-0.06	-.000
PQ Peer Tutoring	12.86	5.27	58	12.24	5.57	41	+0.56	+.057
PQ Discipline	4.85	1.11	103	4.80	0.98	76	+0.32	+.024
PQ Help Neighbor's Child	4.67	2.56	83	4.26	2.73	58	+0.92	+.078
PQ Attit Re Paying Teachers	4.09	1.34	101	3.56	1.52	77	+2.47	+.183
PQ Attit Re Paying Parents	3.82	1.40	103	2.80	1.14	76	+5.17	+.362
PI Help from Adults	10.20	2.35	25	8.92	2.29	25	+1.95	+.271
PI Parent-School Contact	10.68	2.95	25	10.36	2.83	25	+0.39	+.057
PI Parent-Parent Contact	8.12	1.99	25	6.92	1.73	25	+2.28	+.312
PI Satis w/Child's Ach	8.68	1.77	25	8.40	2.33	25	+0.48	+.069
PI Encourage Child	15.24	2.57	25	13.60	2.50	25	+2.29	+.314
PI Student Learning Behav	5.36	1.93	25	5.08	1.29	25	+0.60	+.087

The "Parent Attitude toward Child's Achievement" index at Oakland shows a rise in t-value from -3.29 to -0.45. This shift could be due to a regression effect; however, it occurs entirely through a rise over time in the level of this index at the EXP school. That is, the average at the EXP school changes from 4.89 at Time 1 to 5.29 at Time 2, while the average of the CON school is unchanged at 5.38 for both times. Because two items made up this index, 6 was the logical high score and the CON school might have been constrained by this ceiling effect.

The observed shift in the "Parent Attitude toward Teacher" index at Oakland also occurs primarily through a rise in level at the EXP school. That is, the level for this index at the EXP school changes from 13.03 at Time 1 to 13.66 at Time 2, while the level at the CON school moves only from 13.56 to 13.68.

The last of the Questionnaire-based indexes in Oakland to show a positive effect is the index for "Parent Attitude about Motivation". The t-value on this index shifts from -1.16 at Time 1 to +0.32 at Time 2. The shift is the result of a decline in the preference for praise as a motivational technique in both schools, but a greater decline at the CON school.

The two indexes based on the Questionnaire which show a difference contrary to prediction in Oakland both deal with the payment of incentives. The index "Parent Attitude toward Incentives to Teachers" changes from a t-value of +3.80 at Time 1 to a value of +2.47 at Time 2. This shrinkage is due in part to the fact that although both schools do increase on this index, the CON school exhibits a greater increase in favorableness than does the EXP school. On the second index relating to incentives, "Parent Attitude toward Incentives to Parents", the pattern is similar, although less strong. On this index, part of the explanation for the decline in the t-value is due to the difference in case bases between the Time 1 data and the Time 2 data. It should be noted that the point biserial correlation for this comparison indicates that the effect is quite small, but in the positive direction over time. The positive direction is consistent with the fact that this is one of the sites at which the Parent-Teacher model was used.

The first of the indexes based on the Parent Interview is that for "Adult Help for Students". This index shows a positive shift (from a t-value of -0.53 at Time 1 to a t-value of +1.95 at Time 2). This shift is the result of a large increase in average at the EXP school (from 8.18 to 10.20) and a slight increase at the CON school (from 8.62 to 8.92.). The index for "Parent-Parent Contact" based on the Interview also shows an increase in t-value. This increase also is due to a greater rise at the EXP school than at the CON school. The third Interview index for which a shift large enough to note occurs in Oakland is "Parent Satisfaction with Child's Achievement". The t-value for this index changes from -2.14 to +0.48. This change is due primarily to a rise in the average of the EXP school from 7.50 to 8.68, but a decline in the average at the CON school from 8.65 to 8.40 also contributes. Thus, this pattern could be due to regression effect. Finally, the positive result obtained for the "Parent Supports Child's Learning" index is again due to a pattern in which there is an increase over time at both the EXP and the CON school, but the increase at the EXP school is larger. It should be noted that the composition of this index makes it fairly similar to the index called "Adult Help for Student". Thus, the fact that both indexes show a positive effect is in part due to this similarity.

Overall, the pattern in the Oakland results seems fairly clear. There does seem to be an improvement in the various attitudes and behaviors of the parents in the EXP school in Oakland relative to that of the CON school. Of course, as with any of these findings, the observed result could be due to any number of causes other than an impact of the treatment. However, the consistency of results is strong enough that the general direction, if not each specific result, seems evident.

The last of the individual site comparisons to be examined are those for San Antonio. These results are presented in Table III-89. As in Oakland, they show a generally positive impact of the treatment. Also, as in Oakland, of the 12 indexes based on the Questionnaire, there are four positive shifts large enough to be noted, and only one negative shift of similar size. Three of the positive shifts are on indexes which also

TABLE III-89-1

PARENT QUESTIONNAIRE AND INTERVIEW RESULTS
E/C, SAN ANTONIO, TIME 1

INDEX	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
PQ Help from Adults	12.52	4.30	148	11.82	4.29	82	+1.19	+.079
PQ Parent-School Contacts	2.42	3.65	193	2.64	2.46	107	-0.55	-.032
PQ Parent-Parent Contacts	0.95	1.60	196	1.04	1.17	107	-0.50	-.028
PQ Satis w/Child's Ach	5.20	1.33	199	5.14	1.26	108	+0.4[illegible]	+.022
PQ Care about Child's Ach	6.00	0.00	194	5.97	0.22	107	UNDEF	UNDEF
PQ Satis w/School	6.29	1.86	184	6.93	1.73	101	-2.86	-.168
PQ Satis w/Teacher	13.69	1.87	188	13.91	1.80	104	-1.01	-.059
PQ Peer Tutoring	10.77	5.41	100	11.05	5.24	56	-0.32	-.026
PQ Discipline	4.79	1.11	196	4.74	1.08	107	+0.36	+.020
PQ Help Neighbor's Child	3.26	2.12	143	3.39	2.41	89	-0.45	-.030
PQ Attit Re Paying Teachers	4.29	1.52	194	3.56	1.56	106	+3.99	+.225
PQ Attit Re Paying Parents	3.77	1.59	197	2.98	1.22	107	+4.48	+.249
PI Help from Adults	8.70	2.65	23	9.17	2.76	24	-0.60	-.088
PI Parent-School Contact	10.48	4.27	23	9.42	3.26	24	+0.96	+.142
PI Parent-Parent Contact	6.26	2.75	23	5.21	1.77	24	+1.57	+.228
PI Satis w/Child's Ach	7.57	2.90	23	7.79	1.84	24	-0.32	-.048
PI Encourage Child	13.78	3.09	23	13.71	3.17	24	+0.08	+.010
PI Student Learning Behav	4.30	2.36	23	4.29	2.24	24	+0.02	+.000

TABLE III-89-2

PARENT QUESTIONNAIRE AND INTERVIEW RESULTS
E/C, SAN ANTONIO, TIME 2

INDEX	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
PQ Help from Adults	13.18	4.55	130	12.45	4.35	40	+0.90	+.069
PQ Parent-School Contacts	2.92	3.71	175	3.50	4.90	56	-0.94	-.062
PQ Parent-Parent Contacts	1.48	1.93	174	1.07	0.99	56	+1.51	+.099
PQ Satis w/Child's Ach	5.41	1.25	183	5.09	1.42	55	+1.61	+.104
PQ Care about Child's Ach	6.00	0.00	181	5.96	0.27	55	UNDEF	UNDEF
PQ Satis w/School	6.63	1.83	163	6.55	1.99	53	+0.29	+.020
PQ Satis w/Teacher	13.85	1.82	163	13.26	2.56	54	+1.87	+.126
PQ Peer Tutoring	11.00	5.78	94	11.41	5.14	27	-0.33	-.030
PQ Discipline	4.62	1.03	179	4.70	1.19	53	-0.47	-.030
PQ Help Neighbor's Child	3.81	2.48	147	4.[illegible]4	3.21	47	-2.09	-.149
PQ Attit Re Paying Teachers	4.59	1.48	179	3.86	1.55	56	+3.19	+.204
PQ Attit Re Paying Parents	4.34	1.48	179	3.29	1.47	56	+4.65	+.291
PI Help from Adults	10.00	1.96	24	8.54	2.99	24	+2.00	+.283
PI Parent-School Contact	12.04	3.63	24	12.12	3.35	24	-0.08	-.010
PI Parent-Parent Contact	5.58	2.75	24	4.87	1.68	24	+1.08	+.157
PI Satis w/Child's Ach	7.79	1.96	24	7.96	2.16	24	-0.28	-.041
PI Encourage Child	15.63	2.36	24	14.12	3.43	24	+1.77	+.252
PI Student Learning Behav	5.17	2.12	24	5.17	1.24	24	+0.00	+.000

showed a similar pattern in Oakland. These are: "Parent-Parent Contact", "Parent Attitude toward Child's Achievement", and "Parent Attitude toward Teacher". The other index which shows a positive shift is "Parent Attitude toward School". The Questionnaire-based index which shows a negative shift is the "Adults Help Neighbor's Children" index.

The observed shift in the Questionnaire-based index of "Parent-Parent Contact" occurs through a rise in the average level at the EXP school from Time 1 to Time 2, and an essentially constant level at the CON school. The average at the EXP school on this index is 0.95 at Time 1, and rises to 1.48 at Time 2, while the average at the CON school is 1.04 at the first time and 1.07 at the second time. The shift on the index of "Parent Attitude toward Child's Achievement" also occurs primarily through a rise in the level of the EXP school. On this index, the average among the EXP school parents at Time 1 is 5.20, and this rises to 5.41 at Time 2. The level at Time 1 among the CON school parents is 5.14, and this declines slightly to 5.09. On the index for "Parent Attitude toward School", the parents in the EXP school are lower at Time 1, 6.29 as compared to 6.93 for the CON school. At Time 2, however, the EXP school average is 6.63, and the average at the CON school is lower, only 6.55. The fourth positive result, for the index "Parent Attitude toward Teacher", is another example of a reversal of direction. That is, the EXP school is lower at Time 1, and is higher at Time 2. Most of these patterns do not seem very likely to be the results of any possible regression effect.

The indexes based on the Parent Interview show two positive results, and one result in the direction opposite to that predicted. As was the case in Oakland, the index "Adult Help for Student" and "Parent Supports Child's Learning" both show positive effects. Again, it should be noted th - these two indexes are constructed so that they share some component item., and so this similarity of outcome is not surprising. There is one negative result in San Antonio on the Interview-based indexes. The index "Parent-School Contact" shows a decline at the EXP school relative to the CON school. The pattern on the "Adult Help for Student" index is that the average in the EXP school rises from 8.70 to 10.00, and the

average in the CON school declines from 9.17 to 8.54. On the related index for "Parent Supports Child's Learning", the pattern is that the CON school increases over time, but the EXP school increases more. The negative shift on the "Parent-School Contact" index occurs because there is a rise in the amount of contact at both schools over time, but the rise is greater at the CON school. This pattern seems likely not to be due to regression effects.

The general conclusion for San Antonio from this analysis is that the pattern here in general is similar to that for Oakland. The reported attitudes and behavior of the parents seem to become relatively more favorable in the EXP school than in the CON school over the course of the project.

Model-Based Comparisons--Parent Questionnaire and Interview Results

As with the earlier analyses, the model-based comparisons here are presented primarily as supplementary information for the benefit of readers who wish to see them. However, they should not be interpreted as reflecting very directly the nature of the impact. In view of the considerable diversity between the results of different sites, the model-based comparisons are not clearly interpretable. For the parent results, there is a fairly high degree of consistency between the patterns obtained in Oakland and those obtained in San Antonio. This makes the Parent-Teacher model comparisons somewhat more justifiable, but they still do not add much to the general finding that the two sites both show positive impacts. The comparisons for the Teacher-Only model sites, for the Parent-Teacher model sites, and for the PT-E versus TO-E group, as well as the overall comparison of all EXP schools against all CON schools, are presented on the following pages with a brief discussion of each. Detailed discussion is not provided, because the real interest is on the individual sites.

When examined for the Teacher-Only comparison, of the 12 indexes based on the Parent Questionnaire, four show a negative impact, and three show a positive impact. These results are shown in Table III-90. The four indexes which shift in the direction opposite to the hypothesis are: "Adult Help for Students", "Parent-Teacher Contact", "Parent Attitude toward Child's Achievement", and "Parent Interest in Child's Achievement". The three indexes which show a positive shift are: "Parent Attitude toward School", "Adults Help Neighbor's Children", and "Parent Attitude toward Incentives to Parents". There is only one Parent Interview index showing a shift of note from Time 1 to Time 2, and that is the index for "Parent-School Contact". These results for the Teacher-Only sites are recognizable as a composite of the results for Cincinnati and those for Jacksonville. There seems little to add to what has already been said in the discussion of the sites separately.

The comparison for the Parent-Teacher sites (Oakland and San Antonio) indicates four shifts worthy of note and in the direction predicted, and one shift of note in the direction opposite to what was predicted. The pattern of results is presented in Table III-91 and is similar to that for Oakland and San Antonio individually. There are increases in the Questionnaire-based indexes for "Parent-Parent Contact", "Parent Attitude toward Child's Achievement", "Parent Interest in Child's Achievement", and "Parent Attitude toward School". The only observed outcome contrary to the direction predicted is that there is a shift downward in the index "Parent Attitude toward Incentives to Teachers".

The Interview-based indexes for the PT comparison show three outcomes in the predicted direction which are large enough to note, and no reversals that are noteworthy. The three positive results are for: "Adult Help for Student", "Parent Satisfaction with Child's Achievement", and "Parent Supports Child's Learning". These results again are similar to those already noted when the two sites were discussed individually. In principle, the estimates presented in these tables should provide more accurate information as to what the average effect of the incentives model would be on parent attitudes than would those of either site

(Text resumes on page III-229)

TABLE III-90-1

PARENT QUESTIONNAIRE AND INTERVIEW RESULTS
E/C, TEACHER ONLY, TIME 1

INDEX	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\bar{X}$	S.D.	N	$\bar{X}$	S.D.	N	t-val.	r_{pb}
PQ Help from Adults	13.26	4.15	261	12.62	4.02	165	+1.57	+.076
PQ Parent-School Contacts	2.58	6.52	351	2.23	4.94	236	+0.70	+.028
PQ Parent-Parent Contacts	0.85	2.14	356	0.84	1.41	237	+0.06	+.000
PQ Satis w/Child's Ach	5.21	1.26	364	5.02	1.37	236	+1.76	+.071
PQ Care about Child's Ach	5.99	0.16	362	5.97	0.29	239	+1.08	+.041
PQ Satis w/School	6.34	1.79	353	6.38	1.71	226	-0.27	-.010
PQ Satis w/Teacher	13.23	2.40	343	13.27	2.11	229	-0.18	-.010
PQ Peer Tutoring	12.07	5.35	168	11.25	5.60	110	+1.22	+.073
PQ Discipline	4.87	1.12	361	4.83	1.07	241	+0.48	+.020
PQ Help Neighbor's Child	3.87	2.41	263	3.82	2.51	179	+0.22	+.010
PQ Attit Re Paying Teachers	3.79	1.48	360	3.69	1.48	236	+0.73	+.030
PQ Attit Re Paying Parents	3.09	1.27	362	2.85	1.14	240	+2.38	+.097
PI Help from Adults	6.79	2.60	52	7.35	2.95	43	-0.98	-.101
PI Parent-School Contact	8.35	4.04	52	7.95	3.30	43	+0.51	+.053
PI Parent-Parent Contact	5.67	1.76	52	5.33	2.10	43	+0.88	+.091
PI Satis w/Child's Ach	7.23	2.77	52	7.53	2.81	43	-0.53	-.055
PI Encourage Child	11.13	3.08	52	11.56	3.19	43	-0.66	-.068
PI Student Learning Behav	4.21	1.96	52	4.60	1.61	43	-1.05	-.109

TABLE III-90-2

PARENT QUESTIONNAIRE AND INTERVIEW RESULTS
E/C, TEACHER ONLY, TIME 2

INDEX	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
PQ Help from Adults	12.81	4.26	129	12.89	4.17	111	-0.14	-.010
PQ Parent-School Contacts	2.32	3.64	183	3.05	7.16	149	-1.21	-.066
PQ Parent-Parent Contacts	0.96	1.97	181	1.16	2.45	152	-0.84	-.046
PQ Satis w/Child's Ach	5.11	1.42	198	5.15	1.36	165	-0.23	-.014
PQ Care about Child's Ach	5.97	0.20	198	5.99	0.16	159	-0.91	-.048
PQ Satis w/School	6.82	1.90	181	6.29	1.88	157	+2.57	+.139
PQ Satis w/Teacher	13.51	1.99	187	13.67	2.00	148	-0.71	-.039
PQ Peer Tutoring	11.60	5.45	105	11.26	5.85	81	+0.41	+.030
PQ Discipline	4.89	1.06	195	4.92	1.07	159	-0.27	-.014
PQ Help Neighbor's Child	4.35	2.64	138	3.78	2.40	116	+1.79	+.112
PQ Attit Re Paying Teachers	3.94	1.42	195	3.81	1.51	156	+0.87	+.047
PQ Attit Re Paying Parents	3.32	1.36	196	2.84	1.09	159	+3.60	+.188
PI Help from Adults	6.94	2.42	48	7.05	2.94	43	-0.19	-.020
PI Parent-School Contact	7.80	3.32	48	8.98	3.82	43	-1.58	-.166
PI Parent-Parent Contact	3.90	1.97	48	3.70	1.54	43	+0.53	+.057
PI Satis w/Child's Ach	7.71	3.02	48	7.72	2.33	43	-0.02	-.000
PI Encourage Child	11.19	3.25	48	11.49	3.66	43	-0.42	-.044
PI Student Learning Behav	4.12	1.88	48	4.70	1.81	43	-1.48	-.155

TABLE III-91-1

PARENT QUESTIONNAIRE AND INTERVIEW RESULTS
E/C PARENT TEACHER ONLY, TIME 1

INDEX	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\bar{X}$	S.D.	N	$\bar{X}$	S.D.	N	t-val.	r_{pb}
PQ Help from Adults	13.07	4.20	256	12.78	4.24	201	+0.74	+.035
PQ Parent-School Contacts	2.72	4.02	350	2.31	3.28	261	+1.34	+.054
PQ Parent-Parent Contacts	1.13	1.84	354	1.06	1.73	260	+0.45	+.017
PQ Satis w/Child's Ach	5.06	1.39	359	5.28	1.20	261	-2.02	-.081
PQ Care about Child's Ach	5.98	0.24	358	5.99	0.14	258	-0.47	-.020
PQ Satis w/School	6.07	1.89	339	6.39	1.88	240	-1.99	-.082
PQ Satis w/Teacher	13.39	2.16	342	13.71	1.86	250	-1.88	-.077
PQ Peer Tutoring	11.46	5.45	188	11.43	5.02	143	+0.04	+.000
PQ Discipline	4.87	1.09	358	4.95	1.06	258	-0.97	-.039
PQ Help Neighbor's Child	3.62	2.41	263	3.68	2.42	215	-0.25	-.010
PQ Attit Re Paying Teachers	4.16	1.51	354	3.44	1.48	254	+5.80	+.230
PQ Attit Re Paying Parents	3.68	1.53	357	2.80	1.08	256	+7.80	+.301
PI Help from Adults	8.38	3.06	61	8.88	2.88	50	-0.88	-.084
PI Parent-School Contact	10.57	4.11	61	9.90	3.12	50	+0.96	+.091
PI Parent-Parent Contact	5.90	2.32	61	5.38	1.70	50	+1.32	+.126
PI Satis w/Child's Ach	7.52	2.56	61	8.24	1.79	50	-1.67	-.158
PI Encourage Child	13.33	3.25	61	13.02	3.09	50	+0.51	+.049
PI Student Learning Behav	4.31	1.98	61	4.28	1.86	50	+0.09	+.010

TABLE III-91-2

PARENT QUESTIONNAIRE AND INTERVIEW RESULTS
E/C, PARENT TEACHER ONLY, TIME 2

INDEX	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
PQ Help from Adults	13.31	4.27	209	12.79	4.42	96	+0.98	+.056
PQ Parent-School Contacts	3.08	4.36	277	2.66	5.00	128	+0.85	+.042
PQ Parent-Parent Contacts	1.68	1.96	275	0.78	1.11	130	+3.24	+.159
PQ Satis w/Child's Ach	5.37	1.25	279	5.26	1.29	132	+0.84	+.041
PQ Care about Child's Ach	6.00	0.06	284	5.98	0.17	131	+1.01	+.050
PQ Satis w/School	6.40	1.90	256	6.32	1.97	130	+0.40	+.020
PQ Satis w/Teacher	13.78	1.88	256	13.50	2.24	128	+1.29	+.066
PQ Peer Tutoring	11.71	5.65	152	11.91	5.38	68	-0.25	-.017
PQ Discipline	4.71	1.07	282	4.76	1.07	129	-0.48	-.024
PQ Help Neighbor's Child	4.12	2.54	230	4.48	2.95	105	-1.13	-.062
PQ Attit Re Paying Teachers	4.41	1.45	280	3.68	1.53	133	+4.65	+.224
PQ Attit Re Paying Parents	4.15	1.47	282	3.01	1.31	132	+7.60	+.351
PI Help from Adults	10.10	2.14	49	8.73	2.64	49	+2.82	+.276
PI Parent-School Contact	11.35	3.34	49	11.22	3.19	49	+0.19	+.020
PI Parent-Parent Contact	6.88	2.69	49	5.92	1.98	49	+2.01	+.201
PI Satis w/Child's Ach	8.24	1.90	49	8.18	2.23	49	+0.15	+.014
PI Encourage Child	15.43	2.45	49	13.86	2.97	49	+2.86	+.280
PI Student Learning Behav	5.27	2.01	49	5.12	1.25	49	+0.42	+.044

individually. However, in view of the great amount of site-to-site variation in all these variables, including parent attitudes and reported behavior, it would be presumptuous to use these estimates as if they were particularly accurate. The main point about them is that the pattern of effects in the two sites at which the Parent-Teacher model was used is generally, although not exactly, similar.

The comparison of the EXP schools in the two Parent-Teacher sites with the EXP schools in the two Teacher-Only sites, aimed at indicating something about the marginal impact of the PT model, is presented in Table III-92. This comparison is used in lieu of a more direct one, but it should be remembered that it can be distorted by differences between sites other than those due to the treatment. This comparison indicates a number of favorable shifts on the Questionnaire-based indexes and on the Interview-based indexes. There are 5 shifts in the predicted direction of the Questionnaire-based indexes, and one shift in the direction opposite to that predicted. On the Interview-based indexes, there are again five shifts in the predicted direction, and no shifts in the opposite direction occur.

The last of the model-based comparisons is that for all of the schools, including the four EXP schools on one side and the four CON schools on the other. Because of the dissimilar patterns of impact observed among the TO and the PT schools, this overall comparison can be expected to mask as much as it reveals; and the outcomes observed from it do indeed mask some of the more interesting results. This overall comparison is presented in Table III-93.

Parent Questionnaire and Interview Results--
Discussion

The most interesting result arising from this analysis of the Parent Questionnaire and Parent Interview data is the general pattern of a positive impact of the incentives model at the two sites where the PT model was implemented. These results, while not perfectly consistent from one site to the other, do show enough consistency to justify the conclusion that the academically-related attitudes and behavior of the parents at

(Text resumes on page III-234)

TABLE III-92-1

PARENT QUESTIONNAIRE AND INTERVIEW RESULTS
PARENT TEACHER-E/TEACHER ONLY-E, TIME 1

INDEX	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
PQ Help from Adults	13.07	4.20	256	13.26	4.15	261	-0.53	-.022
PQ Parent-School Contacts	2.72	4.02	350	2.58	6.52	351	+0.35	+.014
PQ Parent-Parent Contacts	1.13	1.84	354	0.85	2.14	356	+1.82	+.069
PQ Satis w/Child's Ach	5.06	1.39	359	5.21	1.26	364	-1.47	-.055
PQ Care about Child's Ach	5.98	0.24	358	5.99	0.16	362	-0.38	-.014
PQ Satis w/School	6.07	1.89	339	6.34	1.79	353	-1.92	-.073
PQ Satis w/Teacher	13.39	2.16	342	13.23	2.40	343	+0.91	+.035
PQ Peer Tutoring	11.46	5.45	188	12.07	5.35	168	-1.07	-.057
PQ Discipline	4.87	1.09	358	4.87	1.12	361	-0.01	-.000
PQ Help Neighbor's Child	3.62	2.41	263	3.87	2.41	263	-1.16	-.051
PQ Attit Re Paying Teachers	4.16	1.51	354	3.79	1.48	360	+3.32	+.124
PQ Attit Re Paying Parents	3.68	1.53	357	3.09	1.27	362	+5.58	+.204
PI Help from Adults	8.38	3.06	61	6.79	2.60	52	+2.94	+.269
PI Parent-School Contact	10.57	4.11	61	8.35	4.04	52	+2.90	+.265
PI Parent-Parent Contact	5.90	2.32	61	5.67	1.76	52	+0.58	+.055
PI Satis w/Child's Ach	7.52	2.56	61	7.23	2.77	52	+0.59	+.056
PI Encourage Child	13.33	3.25	61	11.13	3.08	52	+3.66	+.328
PI Student Learning Behav	4.31	1.98	61	4.21	1.96	52	+0.27	+.024

TABLE III-92-2

PARENT QUESTIONNAIRE AND INTERVIEW RESULTS
PARENT TEACHER-E/TEACHER ONLY-E, TIME 2

INDEX	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	t-val.	r_{pb}
PQ Help from Adults	13.31	4.27	209	12.81	4.26	129	+1.04	+.057
PQ Parent-School Contacts	3.08	4.36	277	2.32	3.64	183	+1.96	+.091
PQ Parent-Parent Contacts	1.38	1.96	275	0.96	1.97	181	+2.24	+.104
PQ Satis w/Child's Ach	5.37	1.25	279	5.11	1.42	198	+2.10	+.096
PQ Care about Child's Ach	6.00	0.06	284	5.97	0.20	198	+2.13	+.097
PQ Satis w/School	6.40	1.90	256	6.82	1.90	181	-2.27	-.108
PQ Satis w/Teacher	13.78	1.88	256	13.51	1.99	187	+1.44	+.069
PQ Peer Tutoring	11.71	5.65	152	11.60	5.45	105	+0.16	+.010
PQ Discipline	4.71	1.07	282	4.89	1.06	195	-1.83	-.084
PQ Help Neighbor's Child	4.12	2.54	230	4.35	2.64	138	-0.81	-.042
PQ Attit Re Paying Teachers	4.41	1.45	280	3.94	1.42	195	+3.46	+.157
PQ Attit Re Paying Parents	4.15	1.47	282	3.32	1.36	196	+6.25	+.275
PI Help from Adults	10.10	2.14	49	6.94	2.42	48	+6.82	+.574
PI Parent-School Contact	11.35	3.34	49	7.79	3.32	48	+5.26	+.475
PI Parent-Parent Contact	6.88	2.69	49	3.90	1.97	48	+6.22	+.538
PI Satis w/Child's Ach	8.24	1.90	49	7.71	3.02	48	+1.05	+.107
PI Encourage Child	15.43	2.54	49	11.19	3.25	48	+7.27	+.598
PI Student Learning Behav	5.27	2.01	49	4.12	1.88	48	+2.89	+.284

TABLE III-93-1

PARENT QUESTIONNAIRE AND INTERVIEW RESULTS
E/C, ALL, TIME 1

INDEX	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\bar{X}$	S.D.	N	$\bar{X}$	S.D.	N	t-val.	r_{pb}
PQ Help from Adults	13.17	4.17	517	12.71	4.14	366	+1.62	+.055
PQ Parent-School Contacts	2.65	5.42	701	2.27	4.15	497	+1.30	+.037
PQ Parent-Parent Contacts	0.99	2.00	710	0.96	1.59	497	+0.30	+.010
PQ Satis w/Child's Ach	5.14	1.33	723	5.15	1.29	497	-0.24	-.000
PQ Care about Child's Ach	5.98	0.20	720	5.98	0.22	497	+0.44	+.014
PQ Satis w/School	6.21	1.84	692	6.38	1.79	466	-1.61	-.047
PQ Satis w/Teacher	13.31	2.28	685	13.50	1.99	479	-1.45	-.042
PQ Peer Tutoring	11.75	5.40	356	11.36	5.27	253	+0.89	+.036
PQ Discipline	4.87	1.10	719	4.89	1.07	499	-0.36	-.010
PQ Help Neighbor's Child	3.75	2.41	526	3.74	2.46	394	+0.03	+.000
PQ Attit Re Paying Teachers	3.97	1.51	714	3.57	1.48	490	+4.61	+.132
PQ Attit Re Paying Parents	3.38	1.44	719	2.82	1.11	496	+7.24	+.203
PI Help from Adults	7.65	2.95	113	8.17	3.00	93	-1.26	-.088
PI Parent-School Contact	9.55	4.21	113	9.00	3.33	93	+1.02	+.071
PI Parent-Parent Contact	5.80	2.08	113	5.35	1.89	93	+1.58	+.110
PI Satis w/Child's Ach.	7.39	2.65	113	7.91	2.33	93	-1.49	-.104
PI Encourage Child	12.32	3.34	113	12.34	3.20	93	-0.06	-.000
PI Student Learning Behav	4.27	1.96	113	4.43	1.75	93	-0.06	-.044

TABLE III-93-2

PARENT QUESTIONNAIRE AND INTERVIEW RESULTS
E/C, ALL, TIME 2

INDEX	RESULTS						STATISTICS	
	EXPERIMENTAL			CONTROL				
	$\bar{X}$	S.D.	N	$\bar{X}$	S.D.	N	t-val.	r_{pb}
PQ Help from Adults	13.12	4.27	338	12.85	4.28	207	+0.73	+.032
PQ Parent-School Contacts	2.78	4.10	460	2.87	6.24	277	-0.26	-.010
PQ Parent-Parent Contacts	1.21	1.97	456	0.99	1.96	282	+1.52	+.056
PQ Satis w/Child's Ach	5.26	1.33	477	5.20	1.33	297	+0.68	+.024
PQ Care about Child's Ach	5.99	0.14	482	5.99	0.17	290	-0.07	-.000
PQ Satis w/School	6.57	1.91	437	6.30	1.92	287	+1.87	+.069
PQ Satis w/Teacher	13.67	1.93	443	13.59	2.11	276	+0.51	+.020
PQ Peer Tutoring	11.67	5.55	257	11.56	5.63	149	+0.19	+.010
PQ Discipline	4.78	1.07	477	4.85	1.07	288	-0.84	-.030
PQ Help Neighbor's Child	4.21	2.58	368	4.11	2.69	221	+0.44	+.017
PQ Attit Re Paying Teachers	4.22	1.45	475	3.75	1.52	289	+4.23	+.151
PQ Attit Re Paying Parents	3.81	1.48	478	2.91	1.20	291	+8.68	+.299
PI Help from Adults	8.54	2.77	97	7.95	2.90	92	+1.43	+.104
PI Parent-School Contact	9.59	3.76	97	10.17	3.66	92	-1.09	-.079
PI Parent-Parent Contact	5.40	2.79	97	4.88	2.10	92	+1.45	+.105
PI Satis w/Child's Ach	7.98	2.52	97	7.97	2.28	92	+0.03	+.000
PI Encourage Child	13.33	3.56	97	12.75	3.50	92	+1.13	+.082
PI Student Learning Behav	4.70	2.02	97	4.92	1.54	92	-0.85	-.062

these two sites was differentially higher in the predicted direction.

These results must, of course, be interpreted with some caution. The relatively low rate of response to the Parent Questionnaire, and the lessening of response rate from Time 1 to Time 2, both make direct generalizations from these results somewhat risky. However, it would be possible to err in the direction of excess caution as well as in the reverse direction. Despite the necessary qualifications, and keeping in mind the various limitations that have already been mentioned about the difficulties encountered in implementing the model and informing parents about the project, these results are definitely grounds to justify further, more precise, investigation.

CHAPTER IV

SUMMARY OF FINDINGS, CONCLUSIONS, AND RECOMMENDATIONS

Findings

Impact Results

Student Achievement Results. The analysis of student achievement was based upon the data from the Metropolitan Achievement Tests, administered in the fall and spring to both experimental and control students. In Reading there was an educationally significant increase in learning rate only in one of the four sites, Oakland. There were negligible increases in two other sites, and an apparent decrease in learning rate at the EXP school in the fourth site, Cincinnati. Thus, the results for Reading seem to be dependent as much on factors that are characteristic of the individual sites as on any general impact of either the TO model or the PT model.

For achievement in Mathematics, there were increases in learning rate in Oakland and San Antonio, which were the two sites at which the Parent Teacher model was implemented. However, these results were due in part to the characteristics of the CON school at each site, and so may reflect the operation of influences quite unrelated to the incentives model. Nevertheless, the mathematics achievement results are suggestive.

Student Attitude and Results. The data obtained from the two instruments used to measure student attitudes show no interpretable pattern. The results from the Student Questionnaire in a number of cases are directly contradictory to those obtained from the Student Interview. This lack of clear pattern is especially evident at Cincinnati and Jacksonville. In Oakland, there are several indexes from the Student Questionnaire which suggests some slight positive impact of the incentives model, but the corresponding indexes from the Student Interview are equally strong in showing trends opposite to the direction predicted. In San Antonio, the results again are not at all clear, although there is a slight

tendency for student self-reported attitudes to decline relatively over the course of the year.

These findings are essentially noninterpretable. A number of possible explanations can be made, including both substantive and methodological factors. However, it must be noted that this set of data does not indicate any positive impact of the incentives model on student attitudes; therefore, to that degree at least, the results do not confirm the general hypothesis.

Student Behavior Results. As with the data from the Student Questionnaire and Student Interview, the data collected from the school records on attendance rates for classroom do not show any discernible trend over time. If there is any such trend, it probably is in a regative direction, because all of the comparisons made indicate that the direction of change is toward a deterioration of the relative position of the EXP school.

The rates of attendance indicated by this data are quite high, probably unusually high for schools serving a low-income, inner-city population. Thus, it may be that the attendance rate could not have been improved by any new school project, because it had already reached a ceiling.

The results obtained from the Classroom Observation of student behavior show no consistent patterns of difference between the EXP and CON schools. Although there are some relative changes in one or another kind of behavior observed from Time 1 to Time 2, these changes turn out to be the result of changes in the CON school as much as in the EXP school. Therefore, it is concluded that they do not reflect any consistent pattern.

Thus, the data gathered on the attitudes and behavior of the students do not reveal any clear impact of the incentives model. The explanations that might be given for this lack of pattern in the data are several. Most obviously, the lack of pattern might reflect the fact that the incentives model made no difference whatever to the students. Alternately, it might be argued that the instruments used, and the circumstances of their administration, led to a lack of sensitivity, so that the true

differences were not picked up. This latter argument loses some force, in view of the fact that there are real inconsistencies and reversals in addition to the absence of positive results in the data. A third explanation is that the initial formulation of the problem was erroneous. It may be that students in the elementary, and particularly the early elementary grades, simply do not respond in a stable way to <u>any</u> new feature of their environment. In a sense, it might be argued that children of this age, as a rule, do not have distinguishable and somewhat consistent attitudes or behavior patterns relative to schooling.

<u>Teacher Attitude Results</u>. As with the data from the Student Questionnaire and Interviews, the results obtained from the Teacher Questionnaires and Interviews are not immediately clear. However, there is considerably more pattern to the results than was found in the Student data.

In Cincinnati, there was no very strong pattern, but there were several changes in the direction opposite to what was predicted. For Cincinnati, it can be concluded that there was a slight negative impact of the incentives model on the attitudes of the teachers involved. The results in Jacksonville include some negative changes also, but there are in addition a number of positive changes indicative of what might be described loosely as classroom "warmth". On balance, it would appear that the impact in Jacksonville, while limited to certain variables, may have been positive. In Oakland, the results for the Teacher Questionnaire and Interview show few changes, and these few changes are in both directions. Thus, there is no real basis for concluding that the incentives model had any specified impact on teacher attitudes in Oakland. Finally, in San Antonio, there were several changes in teacher attitudes over time. One reversal of direction in San Antonio is notable, particularly because San Antonio, from most of the available evidence, seems to have been the site at which the implementation of the project encountered the least opposition among the teachers. The negative change in San Antonio was on the index which measures a teacher's attitude toward modern teaching techniques, such as individualized instruction and

supervised peer tutoring.

One point worth noting is that the teachers' professional associations were the most self-conscious and strongest in Cincinnati, where the impact was somewhat negative, and in Oakland, where there was little trace of any impact. There may well be a political context and social process operating that inhibits the change of individual teacher attitudes.

Teacher Behavior Results. The data on teacher classroom behavior show no clear patterns. In fact, its most striking aspect is the chaotic quality of the differences observed over time. Amid this noise one can only guess at some slight trends. For example, in both Cincinnati and Jacksonville, the changes over time reflect an increase in more traditional and authoritarian behaviors (punishing, criticizing, lecturing), and also increases in some behaviors more characteristic of modern approaches to classroom organization (answering student questions, praising, playing learning games, etc.).

In the Oakland results, the pattern again is very faint. However, if there is a pattern here, it is in the direction of a slight increase in the use of the more traditional behaviors. Finally, in San Antonio, the pattern again is mixed, although with some traces of an increase in traditionalism. It should be noted that there may be specific defects in the data collected on San Antonio classrooms, and so the results from it must be used cautiously. Also, it should be noted that the patterns observed in Oakland and San Antonio, which were the two sites where the Parent Teacher model was used, were in several instances, directly contradictory. That is, in Oakland, a particular behavior would show an increase over time, but in San Antonio, the same behavior would show a decrease over time.

On balance, then, the results of the Teacher Observation data do not point to any clear impacts. Indeed, they tend to vindicate the several teachers whose reaction to the notion of an incentives project was, in effect, that the offer of incentives alone, without any substantive prescriptions, would not cause them to do anything differently because they had been working as hard and as effectively as they were able.

Parent Attitude Results. In contrast to the various results just described, the Parent Questionnaire and Interview data do show some fairly clear patterns. The caution must be made that these patterns could well be due to methodological artifacts, or to real features of the situation at a particular site other than the incentives model. However, the results do appear and should not be ignored.

For Cincinnati, there are few changes, and these are generally in a negative direction. The data for Jacksonville, on the other hand, reflect a larger number of changes over time. These are mostly in a negative direction, and no particular pattern emerges that might account for them.

The results for Oakland and San Antonio, however, as reflected in these two instruments, show a fairly sizable number of changes, almost all of which are in a positive direction. There is one clear exception in both of these sites, however. In both Oakland and San Antonio, there is no increase in the amount of Parent-School Contact observed in this data. Thus, the finding seems to be that the introduction of the incentives model led to an increase in level of parent interest and involvement in the child's education, and perhaps also to a more favorable attitude toward the school. It did not, however, lead to the kinds of joint activities that had been hypothesized. This might be attributable to the delays encountered in delivering the contracts to the parents for their signature. Although there was a great deal of talk about offering incentives to parents, they may have taken a wait-and-see attitude before doing anything with the schools or other parents.

Methodological Results

Introduction. This summary of the methodological issues that emerged in the course of carrying out the implementation of the project, the data collection, and the analysis of the data, will be presented under the following headings: (1) measurement of academic achievement; (2) design features of the project; (3) attitude and behavior measuring instruments; (4) implementation procedures; and (5) general considerations.

Measurement of Academic Achievement. Over the past several years, because there has come to be an increased use of standardized achievement tests as instruments for evaluating the success of various educational programs, issues have emerged about their methodological properties and their limited applicability. In this project, many of these issues were recognized. Some were dealt with at the planning stage, others in the analysis. On balance, however, the issues remain sufficiently unresolved, both in this project and in general, that some caution in interpreting these results must be employed.

The problems include, for example, the possibility that the use of an inappropriate level of a test will cause the obtained score to be influenced by a ceiling or a floor effect. To deal with this possibility, the design of the project implemented what has been called multilevel testing. That is, each child in a classroom was administered the level of test that most closely matched the best estimate of his personal level of Reading achievement. Thus, the ceiling and floor effect should be less serious in this project than in most uses of mass testing. A second set of problems surrounds the choice of score format to be used in analyzing the results. In this project, it was decided to conduct the analysis using not the grade equivalent scores, which have a number of undesirable features commented on by recent writers (*e.g.*, Coleman and Karweit, 1972), but rather to use a score called the "standard score". This score, supplied by the test publisher, is constructed so that it is an equipercentile normed scale score, comparable across all forms and levels and suggested by the publisher as suitable for measuring growth.

Some additional issues emerged and were addressed in the actual analysis. One of these was the possibility of what has been called "fan spread". This means that students who begin with a high initial score tend to grow at a faster rate than students who begin at a lower score. Thus, over a period of time, the distance between the score of a very able child and the score of a less able child tends to increase. This phenomenon complicates analysis considerably and can distort the apparent conclusions. In this project, it was addressed and found to be of relatively small severity.

It is not likely to have created any sizable distortion in the conclusions. However, the whole topic is a new one, and new techniques are being developed by a number of researchers. Thus, this issue may later be a worthy subject for further analysis.

Another technical issue involved in the analysis of the achievement test data was the possibility of what is called regression to the mean. This phenomenon, in which the scores of an initially high scorer tend to drop back toward the average of the group from which he was drawn, operates not only for individual students, but also for social units such as classrooms, grades, and schools. Because of the design of the project, there was no way to determine precisely the severity of regression to the mean. It may well have contributed to some degree to a distortion. This distortion would be such as to make the treatment school appear more successful than it really was. Thus, this distortion would, if anything, contribute to an over-positive interpretation of the results. This fact should be kept in mind when examining the achievement test results. However, the magnitude of the distortion due to the regression effect is unlikely to be large enough to change the general direction of the conclusions or their educational significance.

Two other issues of a methodological nature concern the problems of reducing the amount of noise (in the technical sense) present in the achievement data. In the analysis of this project, no direct attempt was made to reduce this noise, but the desirability of doing so (by means of estimates of reliability, corrections for attenuation, and perhaps the use of predictor variables in some well-defined model) is clear. The limited amount of relevant data and the time pressures involved for completing the analysis precluded that kind of effort for this report.

Design Features of this Project. There were several basic features of the design used for this project which impose limitations on the kinds of inferences which can be drawn. Two of these are central. First, there was no use of the principle of randomized assignment of treatment to units. Second, the number of experimental units (*i.e.*, schools) was quite small.

For these reasons, it must be made clear that the observed differences may be due not to the impact of the incentives treatment but to idiosyncratic differences between individual schools and sites. A third feature of the design was that the four sites which were chosen are relatively heterogeneous. This represents a strategic decision that is inherently difficult. That is, the use of a heterogeneous sample such as was done here allows one to make generalizations more widely if the pattern in the data is clear. However, if the pattern in the data turns out to be not uniform, then this design makes the inference less defensible as to what features of the population may have created those differences.

<u>Attitude and Behavior Measuring Instruments</u>. The instrument used to measure academic achievement was chosen from among several well known and widely used standard achievement tests. In this sense, the choice was noncontroversial, although at a more specific level there are differences between these several tests which make the choice an important one. However, for the instruments used to measure the attitudes and behaviors of the various persons involved in the school, the situation was quite different. For these objectives, there were no readily available instruments which were fully appropriate. Thus, new instruments, in many cases adopting formats from instruments used in other research, were developed. The development aimed at producing a simple, easily administered instrument which would have immediate and direct relevance to the research questions.

A second feature of the instrumentation was an attempt to use the triangulation notion to pin down the actual changes in attitudes and behavior. That is, in several areas, a single broad concept was to be measured by several instruments of different kinds. This approach, desirable as it was, was handicapped by the short time available for development and pretesting of the instruments , and further handicapped by the considerable delay in obtaining OE and OMB forms clearance for the questionnaires. The analysis of the data from these varied instruments indicated a number of unexpected inconsistencies. These inconsistencies create limitations on the confidence that can be placed on any particular interpretation. However, they also indicate that this is an area which deserves further

exploration and attention in future research.

Implementation Procedures. In general, the implementation of the original design can best be described as delayed and partial. The details of these delays will be mentioned in the succeeding section, where some of their implications are used to draw recommendations for future practice. However, at this point it can be noted that the problems in implementation led to two major consequences. First, the treatment simply was not tried at its full strength. That is, the contracts with the participating districts were not signed until after approximately half of the school year had already passed. For some of the parents, the first real awareness of the project came at the time they were asked to fill out the first Parent Questionnaire. This was in March of the school year, after about two-thirds of the year had passed. Thus, whatever impacts are seen here should be regarded as the outcomes of a seriously attenuated treatment. Second, there were delays encountered at two sites because the building principals were unwilling to have the evaluation begin until they thought their schools were prepared.

General Considerations. The general point to be made about the methodology here is simply that the results may well be due to a variety of factors other than the pure impact of a clean incentives model on schools which are otherwise known to be equal. The various methodological problems introduce noise, which makes any patterns difficult to detect, and perhaps creates actual bias. Thus, the data analyzed here should be regarded much as one would regard testimony in a courtroom. The witness is not deliberately trying to create one or another impression, but to tell the facts as he perceived them. Yet, it must be the jury's responsibility in weighing the evidence on data to take into consideration the limitations on the witness's ability to perceive what happened and on the accuracy of his recall. These results, then, like most in social science, must be interpreted as part of a larger context.

Conclusions

The Impact of Incentives

Introduction. Across all participants and variables, a consistent positive trend is not immediately apparent. The influences at work in the four sites and eight buildings--traditions, role relationships, the individual principals--were complex and complicating factors in the analysis of results. However, the analysis did indicate that these and other sources of variance were at least as powerful as the introduction of the incentives models. Therefore, the conclusions offered about the impact of the incentives models are highly tentative.

Single-Site Based Conclusions.

- In Cincinnati, the offer of incentives to teachers, the Teacher Only model, had a negative impact. Student achievement in Reading and Mathematics declined relative to the control school. Teacher attitudes appear to have become more negative. Parent attitudes became very slightly more negative.
- In Jacksonville, the offer of incentives to teachers, the Teacher Only model,had a mixed effect both positive and negative.
- In Oakland, the offer of incentives to parents and teachers, the Parent Teacher model, had a positive impact. Student achievement in Reading and Mathematics increased relative to the control school. Parental attitudes became more positive. However, teacher classroom behavior became slightly more traditional.
- In San Antonio, the offer of incentives to parents and teachers, the Parent Teacher model, had a slight positive effect. Student achievement in Mathematics increased relative to the control school. However, the students' attitudes were slightly more negative. Teacher attitudes were somewhat more positive, but they were also more traditional. Teacher classroom behavior became more traditional. Parental attitudes became more favorable.

Model-Based Conclusions. Although it has been a repetitive theme throughout this analysis, it is important to reaffirm the conviction that the conclusions that follow are based on a very small sample of sites, buildings, and participants. Furthermore, the Teacher Only and Parent Teacher models were not fully or rigorously tested in this demonstration project.

- Incentives to teachers, the Teacher Only model, had an overall negative impact. At one site the impact was markedly negative. At the other site, the impact was both positive and negative.

- Incentives to parents and teachers, the Parent Teacher model, had an overall slightly positive impact. At one site, the results were mostly positive with indications of some slightly negative results. At the other site, there was a slightly positive impact.

The Logistics and Management of Incentives Arrangements

Introduction. In addition to the conclusions reached on the basis of the quantitative analysis of the data, there are some further conclusions to be drawn on the basis of the experiences of this project. Generally, these conclusions deal with the kinds of difficulties which arise in getting an incentives program established and are consistent with the project's objective of demonstrating the feasibility of incentive arrangements in education.

Payment Formulas. The first of the logistical and management problems is one that arises when the payment formula is established. As labor negotiators are well aware, it is often a long step from an agreement in principle about how a payment rate will be determined to an actual set of tables and rules for use in calculating the payment. When incentives are used in education, the problems of establishing a payment formula are increased because of the lack of consensus among the parties involved as to what should be the output upon which they should be measured;

the lack of agreement as to what influences in addition to their own efforts affect the level of output; and finally by disagreements about the technical problems associated with the measurement of academic achievement. In this project, the initial approach was to calculate a single expected growth benchmark for each grade at each school and then to offer payment to any teacher according to the amount by which that teacher's classroom exceeded the benchmark. As a result of objections raised by some teachers, the decision was made to revise this procedure and to develop growth benchmarks for each classroom. Thus, each teacher was working against a more appropriate criterion. The management implication of these events is that even apparently simple and direct incentives models such as these are quite complex when introduced into school settings.

School District Data Management. The second problem that arose in this project has been encountered by other researchers. In general, school district records are not comprehensive or maintained at a useful level of aggregation. Thus, research questions that focus upon the building, grade, or classroom level of aggregation cannot be asked of school district data. Furthermore, those data which are maintained in the buildings are either inaccurate or inconsistently collected and maintained. These limitations of available school documents, both for the previous and current years, reduced the validation that should have been done when unanticipated results were found in this project. The limitations also made the early stage of project implementation, the selection of sites and the choice of school buildings within sites, less objective than it should have been. The conclusion is that unless and until the management of the school districts' data of record is vastly improved, the expectations of, and generalizations from, field projects of this type should be extremely modest.

School District Authority. Although it was not a major difficulty in this project, it should be noted that establishing a full working liaison with the district and defining areas of authority can be very difficult. This is not intended to be a criticism. Rather, it is meant as a reflection upon the functions and structure of school districts and the

researcher's understanding of them. Like many public institutions, school districts, for good reason, are organized to avoid mistakes. Thus, management functions and authority are shared by many people within the administrative hierarchy. No one person has an unchallengeable discretionary control over the kinds of resources that must be quickly mobilized if a project of this sort is to be implemented. For example, the Cincinnati board overruled its management on the type of incentives model to be used. And, in Oakland, the original control school faculty withdrew from the project at the last minute. In effect, management was overruled twice in this project, once by the "board of directors" and once by the "workers".

Furthermore, because authority is diffuse and subject to numerous checks and balances within school districts, the fact that an agreement or contract has been signed is no guarantee that the evaluation plan will be accepted. Thus, in Oakland and Cincinnati, building principals refused to allow data gathering to be initiated until they felt that their schools were prepared for it. One of the anomalies of power relationships in school districts is that, despite appearances, the building principals and their faculties have the potential for a great deal of autonomy. It only appears that they are locked into a position of powerlessness because the school buildings have such limited opportunities to make choices. When the opportunity arises, it is surprising how quickly this latent power emerges.

The conclusion to be drawn is that two months is not enough time, and the approval of only the central office is not sufficient authority, to install a field project in a school district. Furthermore, the willing consent of the building principal and the faculty is the only assurance that the evaluation, so often a source of irritation to people who are attempting to get on with the job, will be implemented as planned once the project has been installed.

Parents and Schools. This project, at least in the two districts that had the Parent Teacher incentives model, necessitated communications with the parents that were different in kind and degree from the communication procedures normally practiced by school districts. At a minimum,

it was required that all parents be informed of the project's objectives and procedures. Ideally, a sustained and continual dialogue between the parents and the schools would have emerged as a desirable outcome.

That the minimum requirement was barely satisfied and that the ideal communications pattern was not realized should not be a surprise. Schools are required by law and organized to teach, socialize and care for children. Parents, despite some populist suspicion of schools, are generally satisfied that these outcomes are being achieved. Thus, the existing communications system between homes and schools consists of a ritualistically attended "open house" and report cards. This is a perfectly functional relationship, supported by a lengthy tradition. Consequently, the implementation of a project which requires the mobilization of parents will strain the existing system beyond its potential if additional resources and time are not devoted to this problem area.

OE Project Administration. Two administrative problems within OE seriously impeded the implementation and evaluation of this project. First, the grants to the districts were delayed as they passed through the required internal approval process. Second, the questionnaires were inordinately delayed in the OE-OMB clearance process. The time constraints of this project were not very different from those of similar past and, presumably , future field efforts; therefore, these delays deserve some attention. It is an axiom of organizational theory that authority must be commensurate with responsibility. In this project, this relationship did not exist. The OE Project Officer found that he had the power only to persuade or cajole others to free the grants to the districts because of insufficient authority. In the process of securing the grants, his project was delayed and the morale of the group that released the funds was not improved because the concept of incentives was not among their approved fiscal year objectives. The necessity for the review and approval of the questionnaires is acknowledged. However, the four to five month clearance process encountered in this project is questionable when the instruments or procedures are not materially altered.

Because of administrative delays, the impact of the incentives models was probably attenuated and the evaluation design was certainly weakened. One can conclude that field experimentation of this type requires a significant reorientation of the priorities and consolidation of authority within the agency responsible for the evaluation.

Recommendations

Discussion

If the idea of incentives has any merit in public education, then the limitations of this project should be discussed before the recommendations are presented. First, the incentive models were intentionally narrow in their scope. Incentive plans can differ along four dimensions: the recipients, the timing, the criteria, and the type of incentive. This project offered incentives only to teachers or parents and teachers; important other participants in the learning process such as principals and students were not included. Along the second dimension, timing, the incentive awards were not to be given until the end of the year, rather than offering interim progress payments. The criteria for payment were classroom reading and mathematics standardized achievement test scores; other units of aggregation such as the grade level could have been used, along with many other outcomes and measurement instruments. Finally, the incentives were cash payments which are extrinsic rewards (and essentially demeaning according to the statements of some teacher groups).

The second limitation is that the incentive models received less than a full exercise and evaluation in this project. In retrospect, the implementation fell short in three areas. The school districts were not recruited into the project until the middle of the summer; hence the buildings were not selected with the consent of the teachers and parents who would participate. Second, the grants to the districts were delayed; therefore, the parents and teachers did not have contracts to sign until half way through the school year. Finally, administrative delays prevented several of the evaluation instruments from being administered at the start of the project. Among the remaining instruments that were not administratively delayed, with the exception of the achievement tests, several were delayed by the building principals in two sites.

Finally, when considering these recommendations, it should be remembered that the idea of incentives arouses latent but very powerful feelings

that are central to the American ethos. The idea is more than just controversial on technical grounds: it is subject to political exploitation. For this reason, the experimentation with, and advocacy of, incentives in education must be extremely cautious. Regardless of the results of this and other incentives projects, it should be remembered that incentives are not likely to be a panacea in education, any more than they have been in industry. Yet, they may well be, as in industry, a useful technique in certain forms and circumstances. The central point about incentives is that they are capable of great harm as well as benefit; therefore, they should be used with care.

Recommendations

Incentives in Education.

1. Based upon the evaluation of this project and the evaluation of the OEO teacher incentives projects in Stockton, California, and Mesa, Arizona, the Teacher Only model should not be further researched or advocated.

2. Based upon the results of the evaluation of the Parent Teacher model, it is recommended that field research of this model be implemented.

3. Based upon this project and some knowledge of the OEO performance contracting experiment, it is recommended that any future field research into incentives in education recognize that incentives are an inextricable part of the accountability movement in public education. Consequently, incentives implicitly involve politics and changes in the distribution of control within school buildings and the school district. The minimal elements of such a recognition in either field experimentation or advocacy of incentives would be:

    o Parents, teachers and others who are the participants must have the opportunity to choose freely to participate.

- The participants must have the opportunity to participate actively in the design of the incentives projects, including such areas as the types of rewards, their timing, etc.
- The participants must have a degree of control over their working environment that is commensurate with the degree to which they are being held accountable.
- The participants must be provided the resources that they feel are necessary for them to meet their commitments.
- The school building is the accountable unit and all professionals in that unit--teachers, the principal, supervisors--share the responsibility.
- The criteria by which performance is measured and rewarded should reflect the range of student outcomes for which the school personnel are responsible.

<u>The Conduct of Field Projects</u>.

4. At a minimum, field projects should be operated on a three-phase basis, with at least one year for each phase. Phase 1 would be devoted to recruitment, implementation, instrument development, design refinement, and the collection of baseline data. Phase 2, the second academic year, would involve the actual operation of the project and the collection of the primary evaluation data. Phase 3 would include the analysis of the primary data and the conduct of follow-up studies.

5. The agency that designs and initiates a field project and the agency that implements the project, generally, the school district, should both participate in funding the project. Past and present field projects have been funded exclusively by the Federal Government, with the districts providing funds or services that are only opportunity costs. The districts do not have a direct fiscal incentive to see that the projects are operated smoothly or evaluated fully. One way to gain a responsible

commitment from the cooperating districts would be to require that the school districts fund a portion of the costs of the project.

6. If the Office of Education is going to continue to operate field endeavors that are based upon a nationwide experimental design and that require a complex evaluation plan, the authority of the officials responsible for the project must be extended to all areas of the project's operation. Furthermore, the responsiveness of middle level OE officials to such research efforts must be addressed and their cooperation guaranteed.

CHAPTER V

REFERENCES

Battelle Columbus Laboratories. Final report on the Office of Economic Opportunity experiment in educational performance contracting. Columbus: Author, 1972.

Bereiter, C. Some persisting delimmas in the measurement of change. In C.W. Harris (Ed.), *Problems in measuring change*. Madison, Wisc.: University of Wisconsin Press, 1963.

Campbell, D.T. Temporal changes in treatment-effect correlations: A quasi-experimental model for institutions, records and longitudinal studies. *Proceedings of the 1970 Invitational Conference on Testing Problems*. Princeton, N.J.: Educational Testing Service, 1971.

Campbell, D.T. and Erlebacher, A. How regression artifacts in quasi-experimental evaluations can mistakenly make compensatory education look harmful. In J. Hellmuth (Ed.), *Compensatory education: A national debate. Vol. 111, The disadvantaged child.* New York: Brunner/Mazel, 1970.

Carpenter, P. and Hall, G.R. Case studies in educational performance contracting, conclusions and implications. Report No. R-900/1-HEW. The RAND Corporation, Santa Monica, California, 1971.

Coleman, J.D. *Introduction to mathematical sociology*. Glencoe, Ill.: The Free Press, 1964.

Coleman, J.S. The mathematical study of change. In H.M. Blalock, Jr. and A.B. Blalock (Eds.), *Methodology in social research*. New York: McGraw-Hill, 1968.

Coleman J.S. and Karweit, N.L. *Information systems and performance measures in schools*. Englewood Cliffs, N.J.: Educational Technology Publications, 1972.

Cronbach, L.J. and Furby, L. How should we measure "change"--or should we? *Psychological Bulletin*, 1970, *74*, 68-80.

Education Turnkey Systems, Inc. Final report from the monitoring and assistance contractor of the use of incentives project. Contract No. OEC-0-71-4768), Washington, D.C., 1972a.

Educational Turnkey Systems, Inc. Final evaluation report, use of incentives project. Contract No. OEC-0-71-4768), Washington, D.C., 1972b.

Hall, G.R., Carpenter, P., Haggart, S.A., Rapp, M.L. and Summer, G.C. A guide to educational performance contracting. Report No. R-955/1-HEW. The RAND Corporation, Santa Monica, California, 1972.

Harcourt Brace Jovanovich, Inc. Guidelines for standardization sampling. Metropolitan Achievement Tests Special Report, 1971; Report #7.

Harcourt Brace Jovanovich, Inc. Means, medians, and standard deviations in terms of raw scores, standard scores, and grade equivalents for national standardization groups. Metropolitan Achievement Tests Special Report, 1971, Report #6.

Harcourt Brace Jovanovich, Inc. Means, medians, and standard deviations in terms of raw scores, standard scores, and grade equivalents for national standardization groups. Metropolitan Achievement Tests Special Report, 1971, Report #8.

Harcourt Brace Jovanovich, Inc. Reliability estimates and standard error of measurements. Metropolitan Achievement Tests Special Report, 1971, Report #10.

Harcourt Brace Jovanovich, Inc. Intercorrelations among subtests. Metropolitan Achievement Tests Special Report, 1971, Report #11.

Harris, C.W. (Ed.) *Problems in measuring change*. Madison, Wisc.: University of Wisconsin Press, 1963.

Jung, S.M., Lipe, D. and Canter, S. Experimental assessment of an incentive program to enhance school learning: A pilot study. Final Report No. AIR-2054-7/71-FR, Palo Alto: American Institutes for Research, 1971.

Lipe, D. and Jung, S.M. Manipulating incentives to enhance school learning. *Review of educational research*, 1971, *41*, 249-280.

Lord, F.M. The measurement of growth. *Educational and psychological measurement*, 1956, *16*, 492-437.

Lord, F.M. Further problems in the measurement of growth. *Educational and psychological measurement*, 1958, *18*, 437-451.

Lord, F.M. A paradox in the interpretation of group comparisons. *Psychological bulletin*, 1967, *68*, 304-305.

Rosenthal, R. *Experimenter effects in behavioral research.* New York: Appleton, 1966.

Saretsky, G. The OEO P.C. experiment and the John Henry effect. *Phi Delta Kappan*, May, 1972, 579-581.

Science Research Associates. Evaluating educational growth with the SRA achievement series and the Iowa test of educational development. Chicago, 1967.

Stucker, J.P. and Hall, G.R. The performance contracting concept in education. Report No. R-699/1-HEW, The RAND Corporation, Santa Monica, California, 1971.

Werts, C.E. and Linn, R.L. A general linear model of studying growth. *Psychological bulletin*, 1970, *73*, 17-22.

Westinghouse Learning Corporation and Ohio University. The impact of Headstart. Vol. I. Report to the Office of Economic Opportunity. 1966.

Winer, B.J. *Statistical principles in experimental design*. New York: McGraw-Hill, 1962.

APPENDIX A
INSTRUMENTS

QUESTIONNAIRES

Student

Teacher

Parent

INTERVIEWS

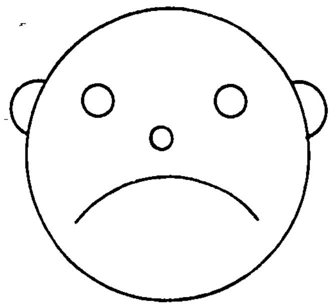
Student

Teacher

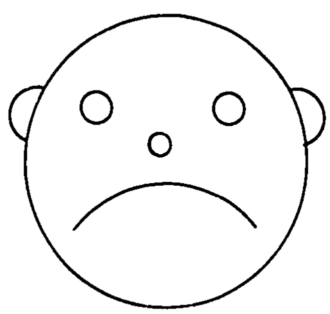
Parent

INCENTIVES IN EDUCATION
STUDENT QUESTIONNAIRE

(NOT IN ORIGINAL FORMAT)

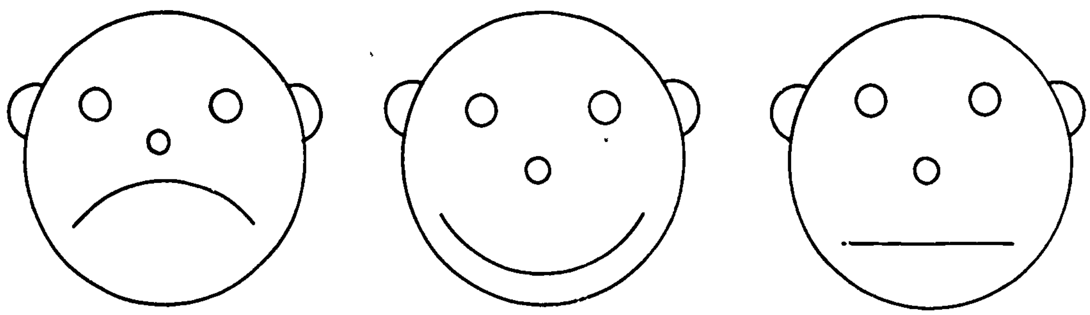


1. When I think about learning to read, I feel like:

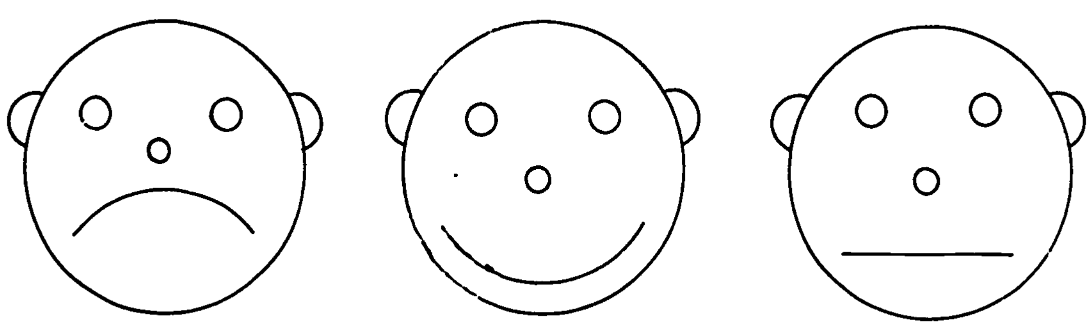


2. When I think about learning to do arithmetic problems, I feel like:

INCENTIVES IN EDUCATION
STUDENT QUESTIONNAIRE

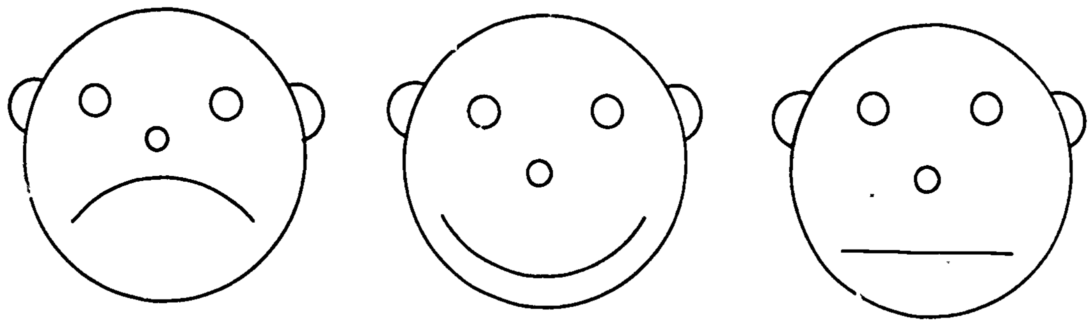


3. When I think about reading, I feel like:

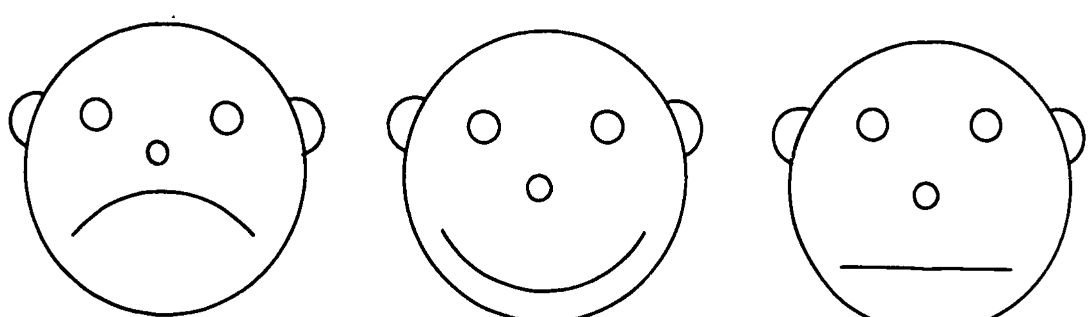


4. When I think about doing arithmetic problems, I feel like:

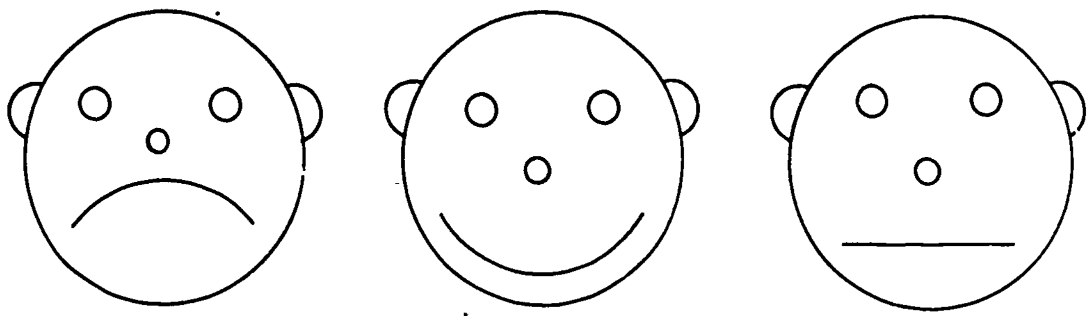
INCENTIVES IN EDUCATION
STUDENT QUESTIONNAIRE



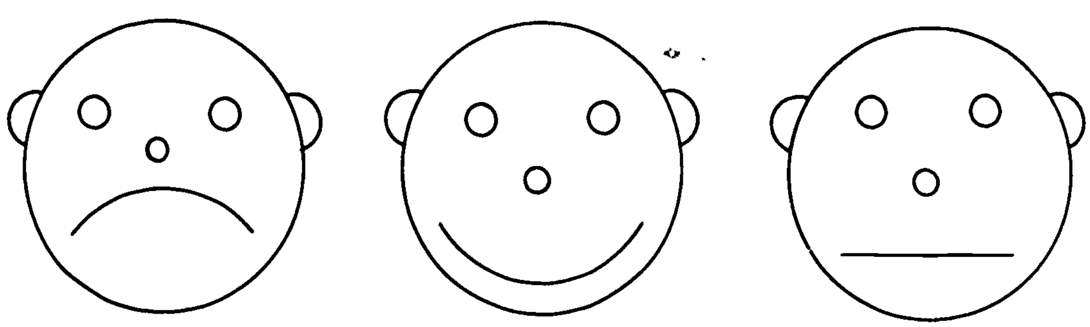
5. When I think of my reading teacher, I feel like:



6. When I think of my arithmetic teacher, I feel like:

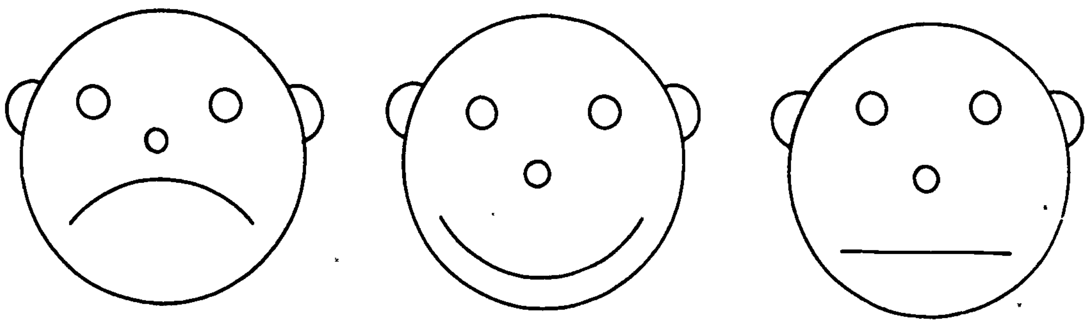


7. When I think about learning other subjects, I feel like:

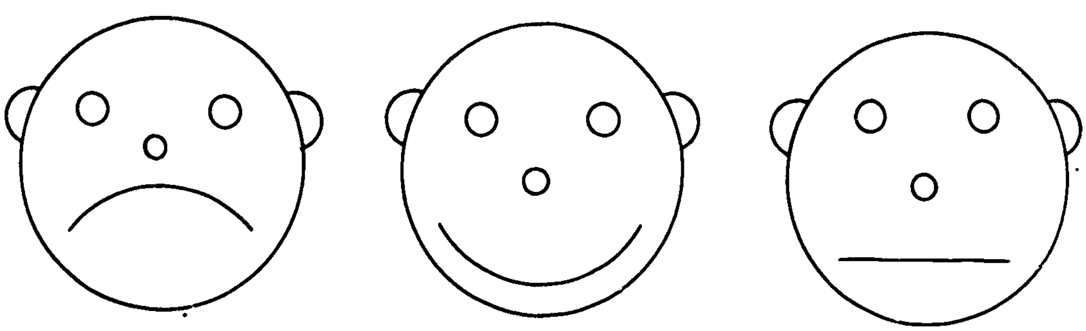


8. When I talk to my teacher in school the teacher looks like:

INCENTIVES IN EDUCATION
STUDENT QUESTIONNAIRE

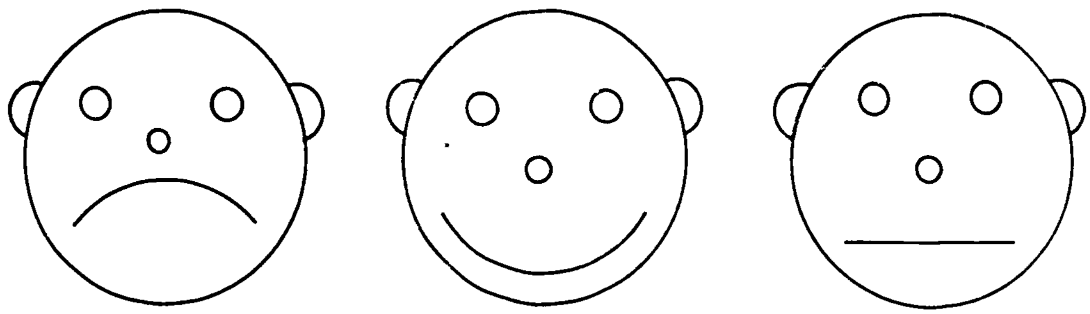


9. When other kids ask me for help in their school work, I feel like:

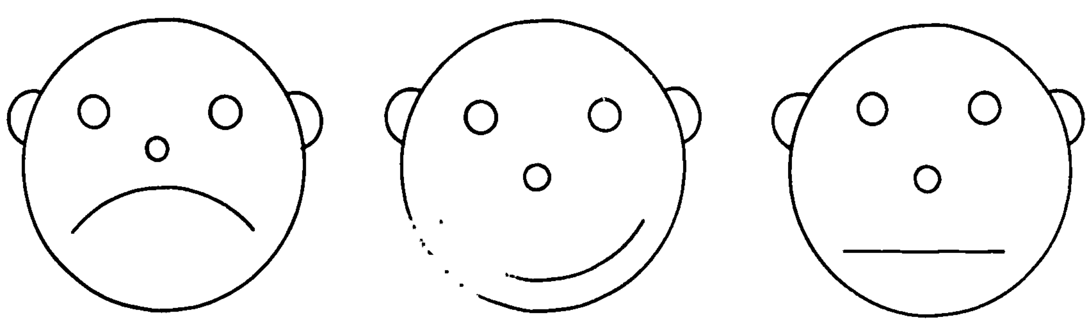


10. When one of the kids asks a question, the teacher looks like:

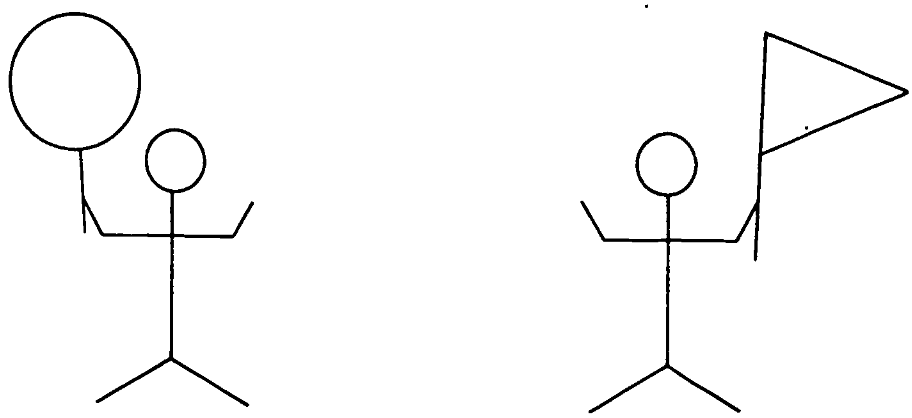
INCENTIVES IN EDUCATION
STUDENT QUESTIONNAIRE



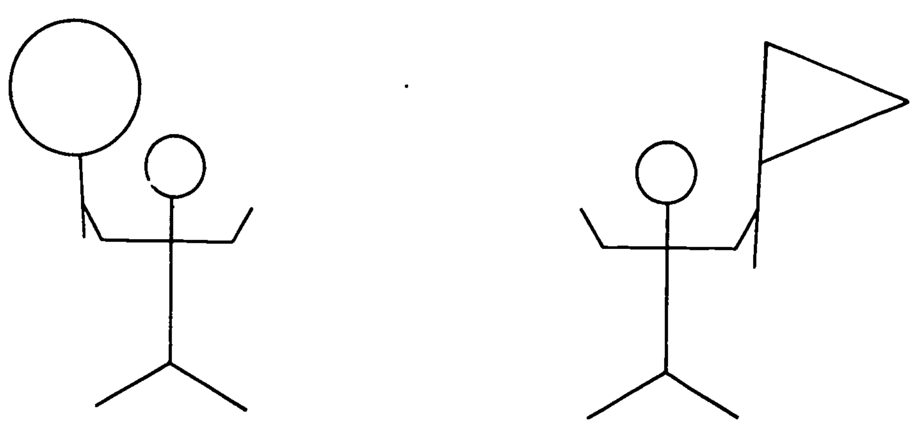
11. When another kid in my class makes a mistake, I feel like:



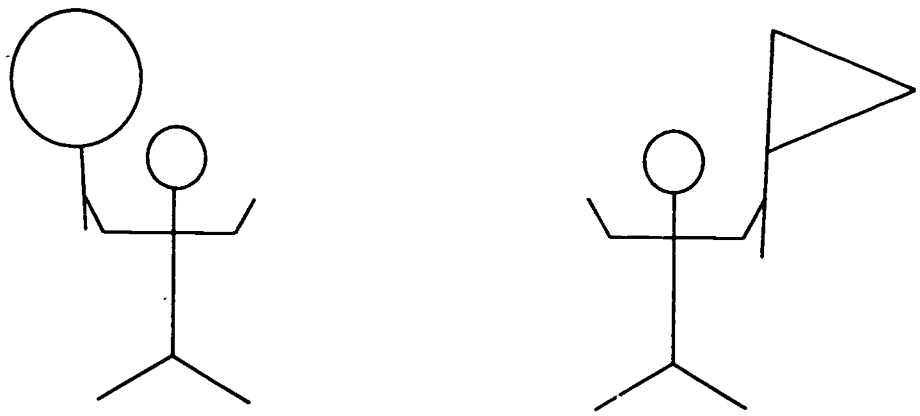
12. When we try to have fun in class, our teacher looks like:



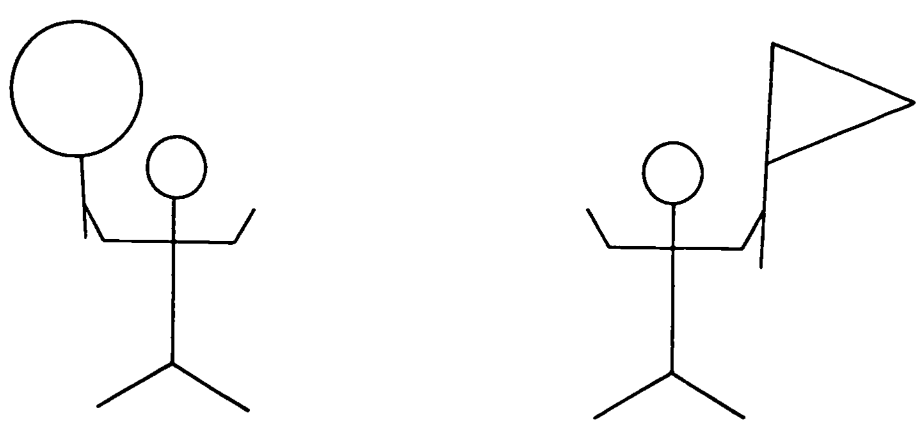
13. The balloon-child goes to a good school; the flag-child goes to a bad school.



14. The balloon-child is not good at doing arithmetic problems; the flag-child is good at doing arithmetic problems.

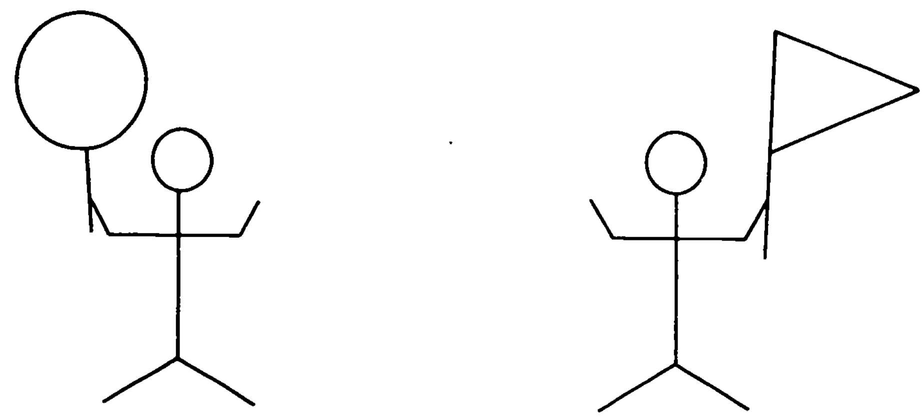


15. The balloon-child does not read well; the flag-child reads well.

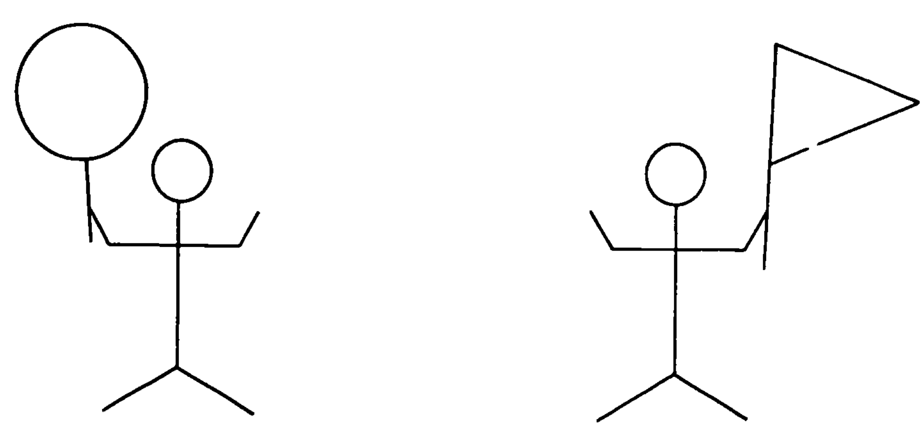


16. The balloon-child likes to go to school; the flag-child does not like to go to school.

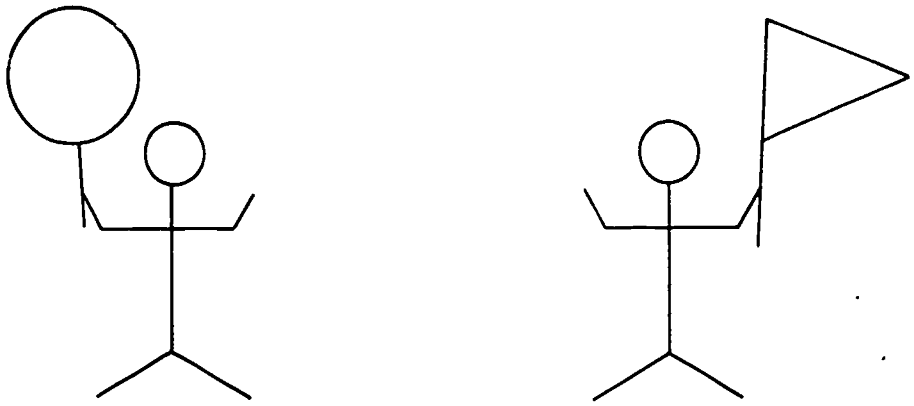
INCENTIVES IN EDUCATION
STUDENT QUESTIONNAIRE



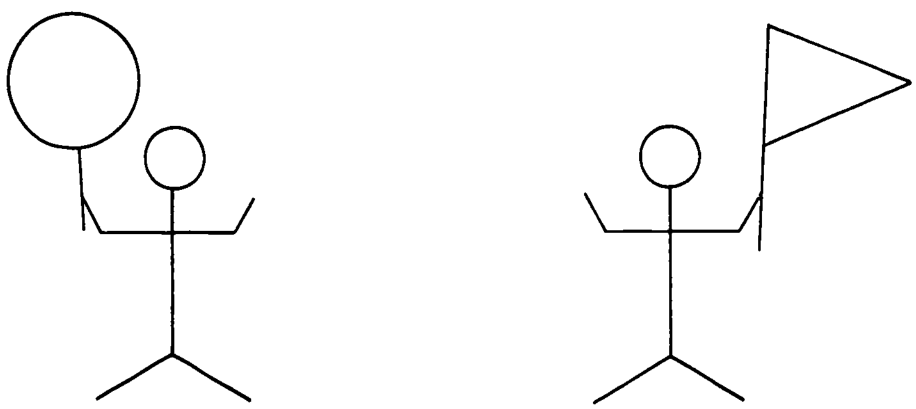
17. The balloon-child likes to work at school with arithmetic problems; the flag-child does not like to work at school with arithmetic problems.



18. The flag-child likes to read at school; the balloon-child does not like to read at school.

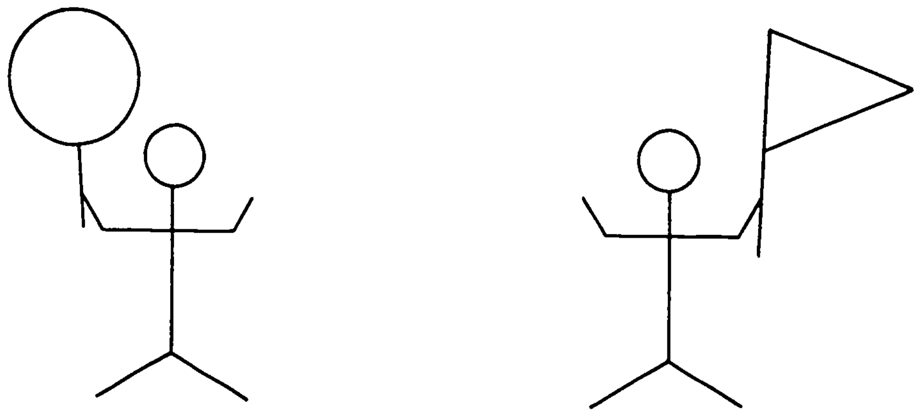


19. The flag-child reads comic books; the balloon-child does not.

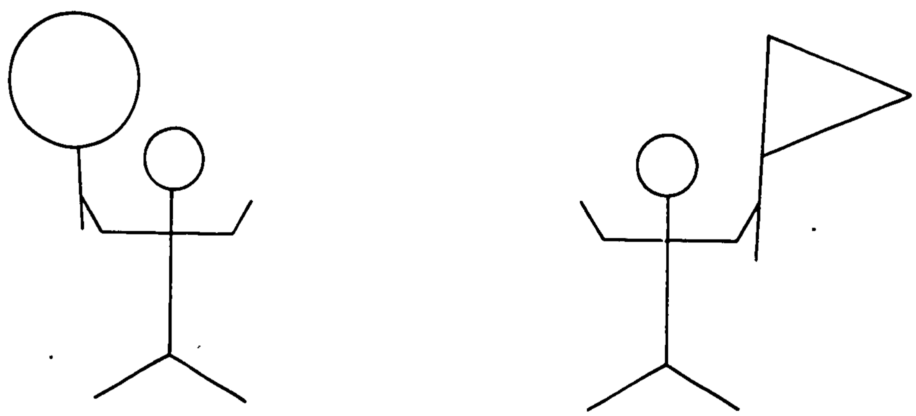


20. The balloon-child reads a lot at home for fun; the flag-child watches TV or plays.

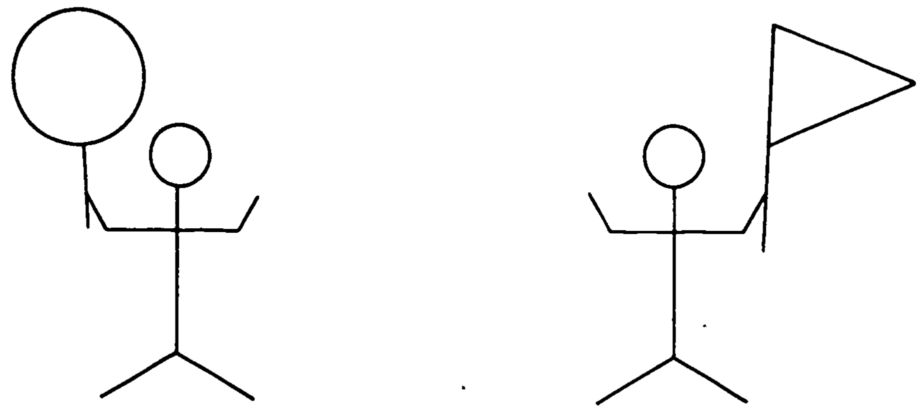
INCENTIVES IN EDUCATION
STUDENT QUESTIONNAIRE



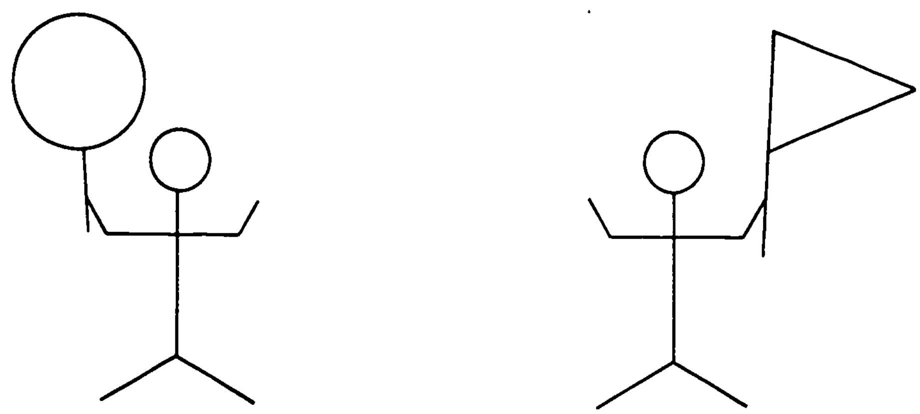
21. The flag-child goes to a library to read for fun; the balloon-child does not go to a library to read for fun.



22. The flag-child does arithmetic problems at home for fun; the balloon-child watches TV or plays.

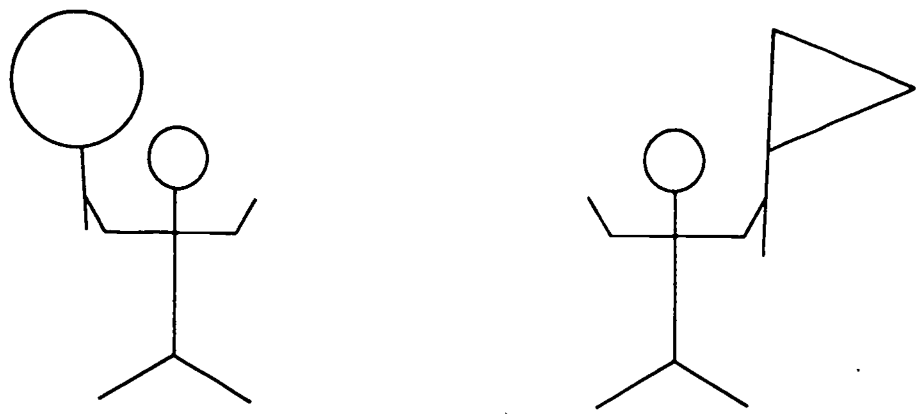


23. The balloon-child writes notes to other people; the flag-child does not write notes to other people.

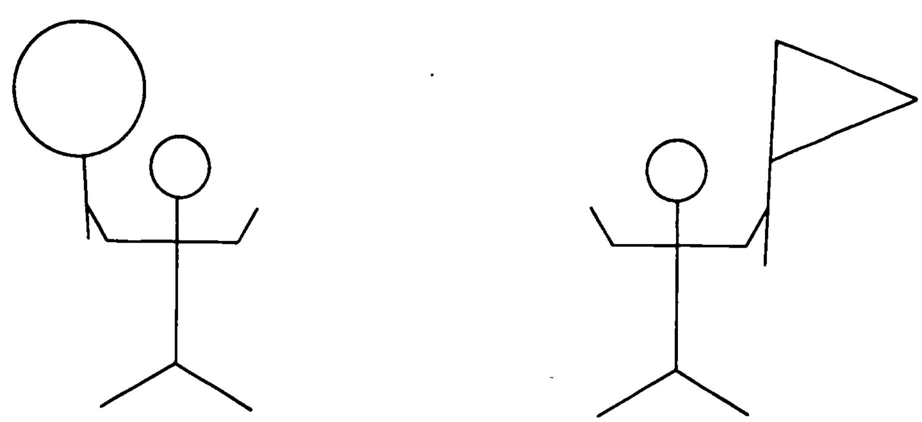


24. The flag-child helps other kids outside of school with their schoolwork; the balloon-child does not help other kids outside of school with their schoolwork.

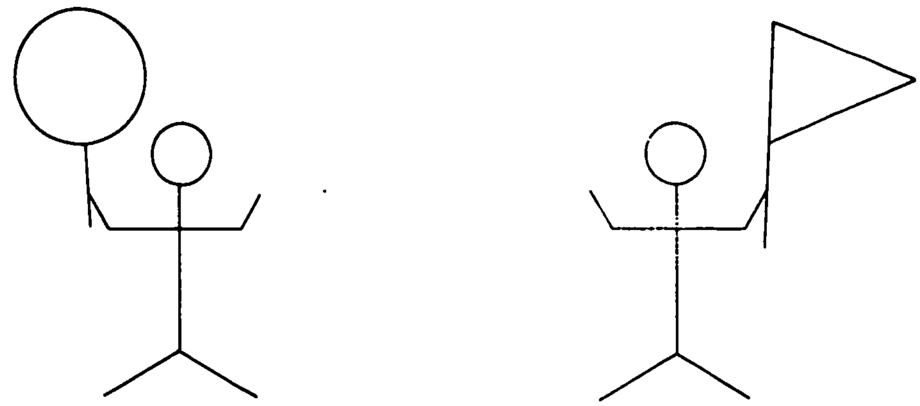
INCENTIVES IN EDUCATION
STUDENT QUESTIONNAIRE



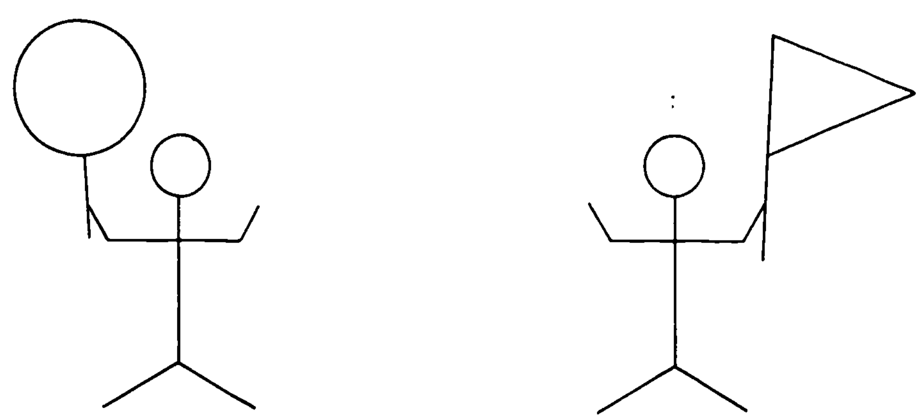
25. The balloon-child helps other kids at school with their schoolwork; the flag-child does not help other kids at school with their schoolwork.



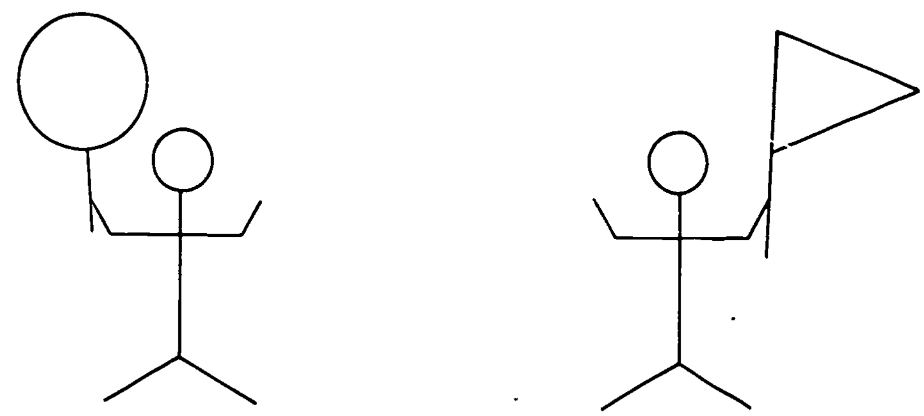
26. The balloon-child wants all the kids to do well in school; the flag-child does not care how the kids do.



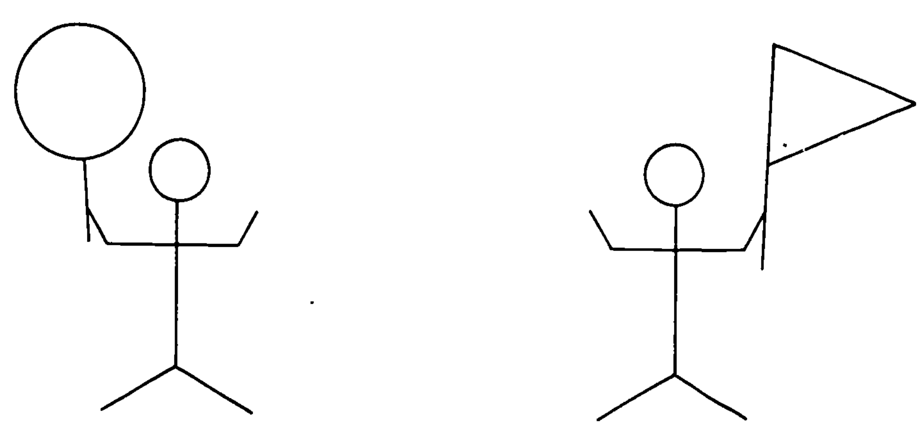
27. The flag-child wants to do well in school; the balloon-child does not care how well he does in school.



28. The balloon-child wants to be good at doing arithmetic problems; the flag-child does not care.

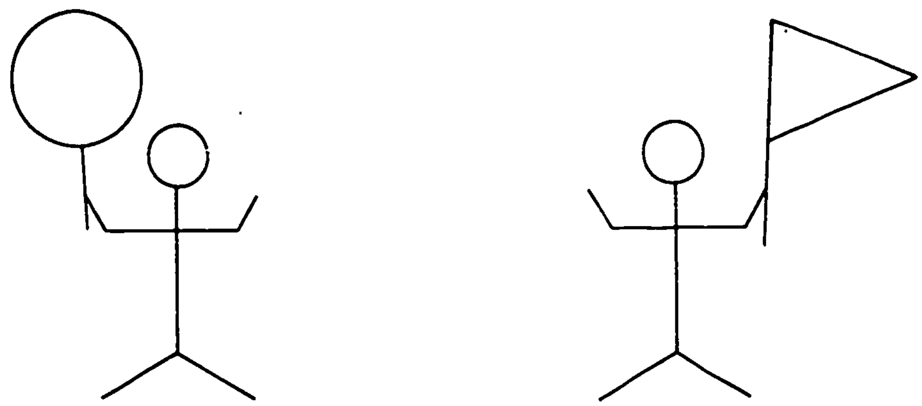


29. The balloon-child wants to be a good reader; the flag-child does not care.

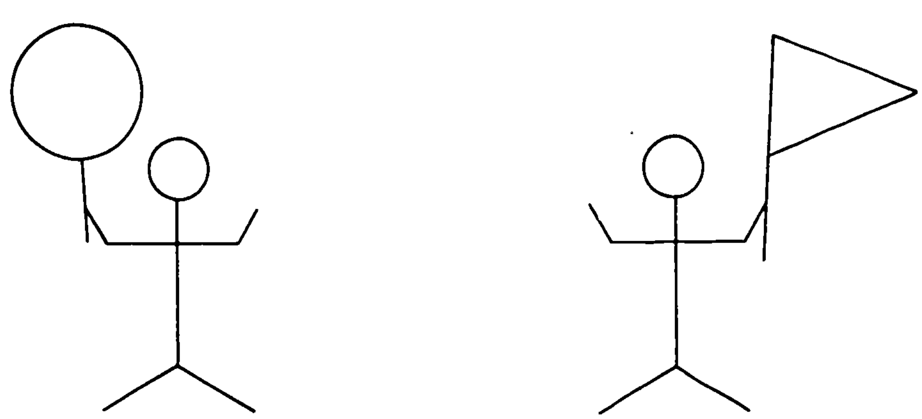


30. Most people think the flag-child is good. Most people think the balloon-child is bad.

INCENTIVES IN EDUCATION
STUDENT QUESTIONNAIRE

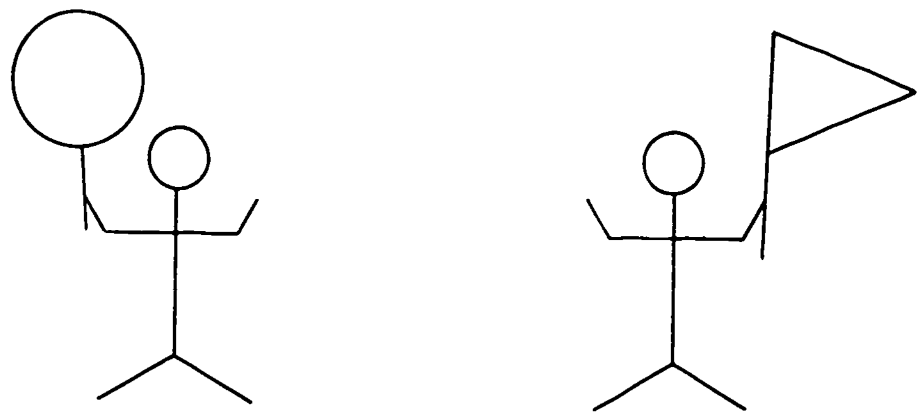


31. The balloon-child expects to be a good reader; the flag-child does not expect to be a good reader.

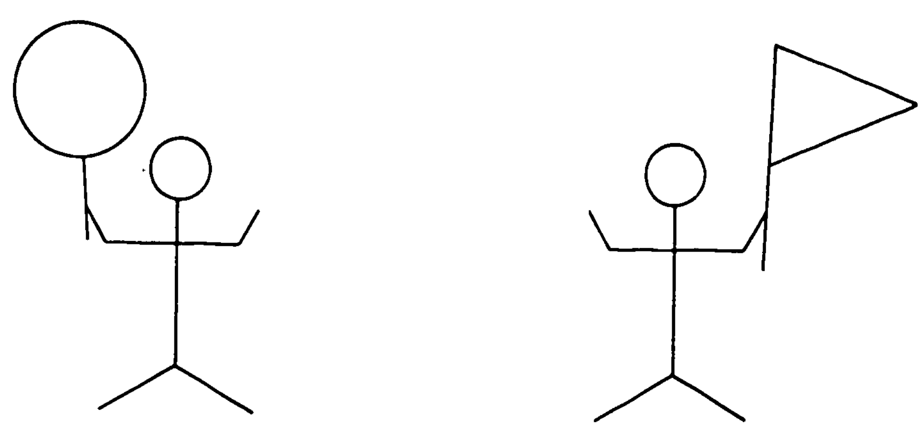


32. It is hard for the balloon-child to learn things; it is easy for the flag-child to learn things.

INCENTIVES IN EDUCATION
STUDENT QUESTIONNAIRE

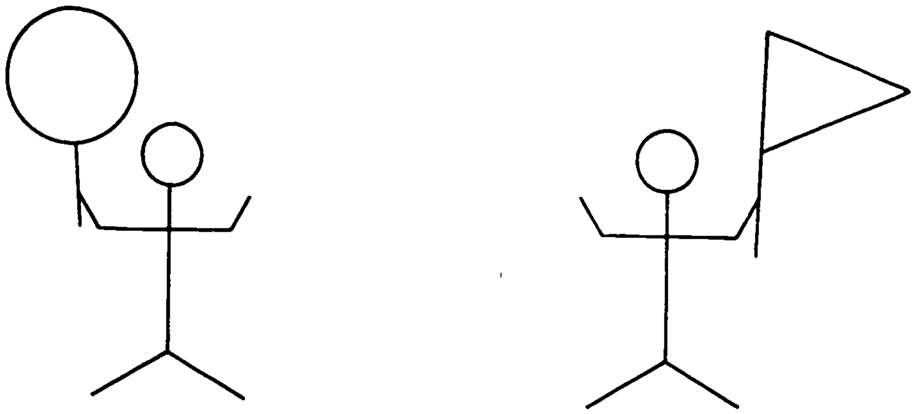


33. The balloon-child expects to go to high school; the flag-child does not expect to go to high school.

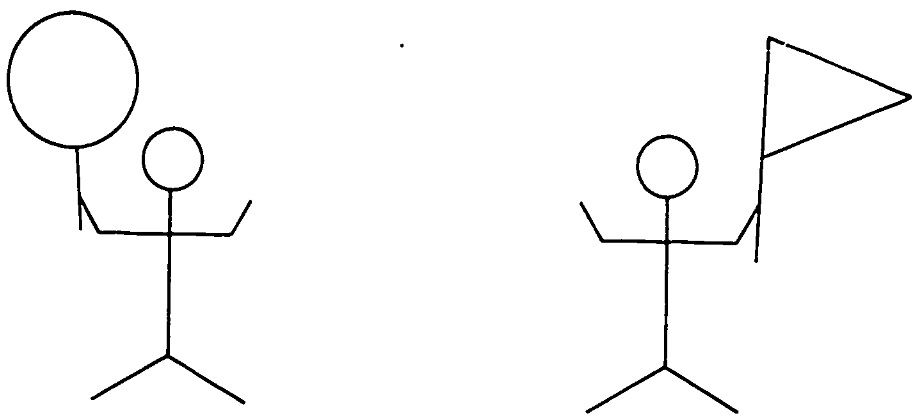


34. The flag-child can do the things he wants to; the balloon-child cannot do the things he wants to.

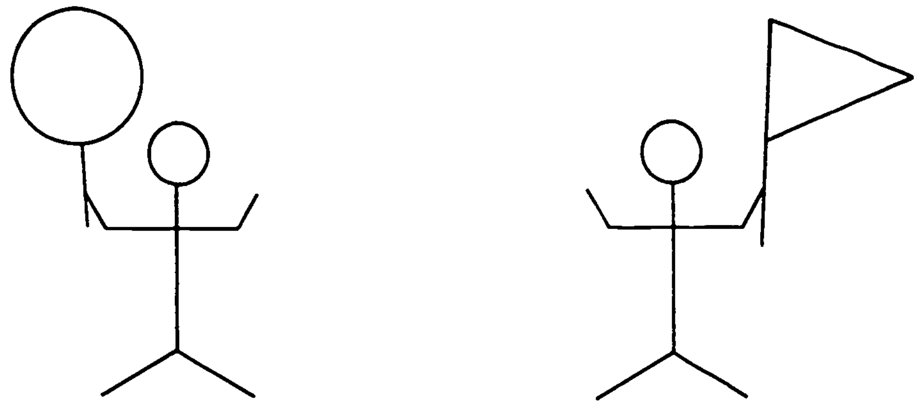
INCENTIVES IN EDUCATION
STUDENT QUESTIONNAIRE



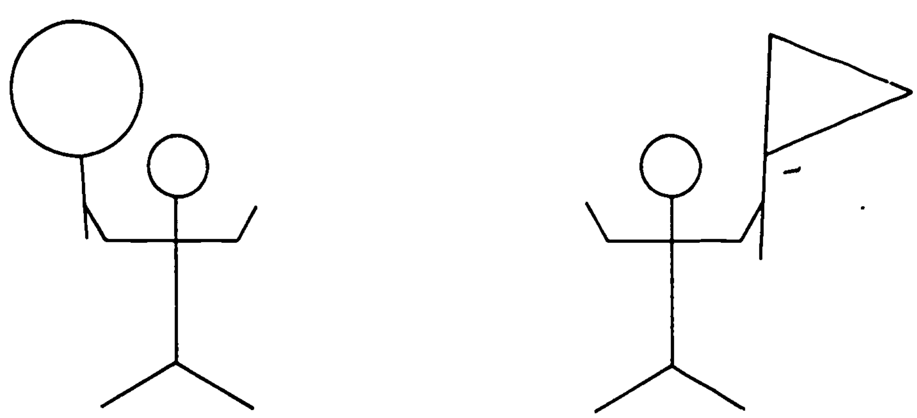
35. The balloon-child expects to be good in arithmetic; the flag-child does not expect to be good in arithmetic.



36. The balloon-child gives up easily on his work; the flag-child finishes his work.

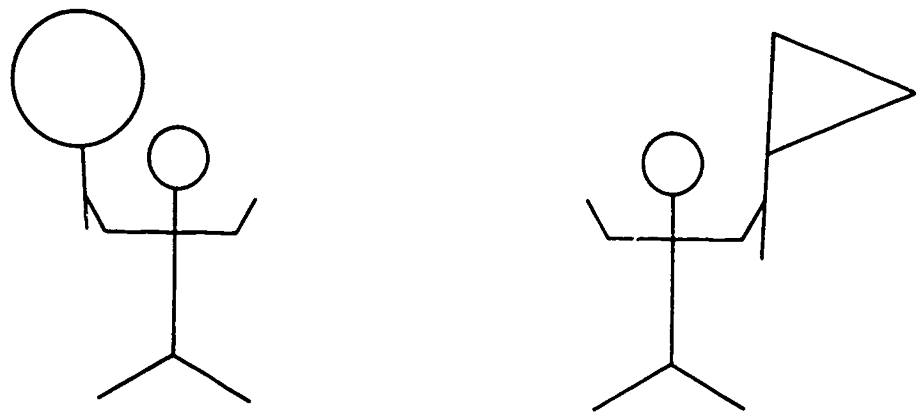


37. The balloon-child does not like himself most of the time; the flag-child likes himself most of the time.

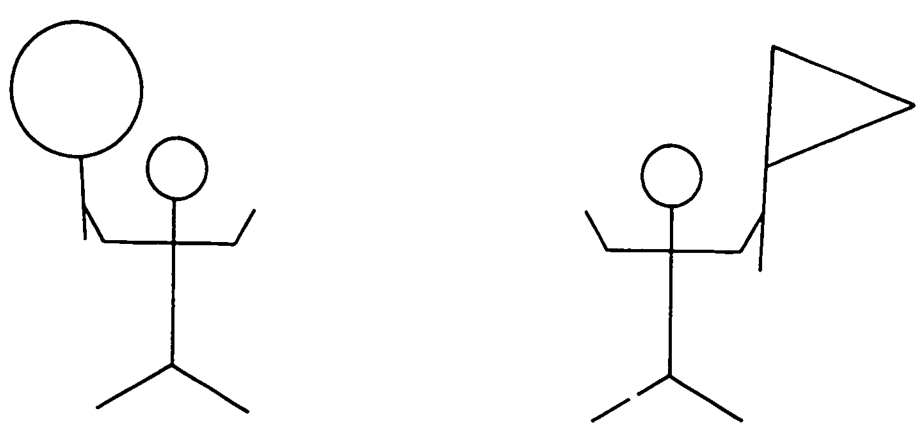


38. The balloon-child wants to go to high school; the flag-child does not want to go to high school.

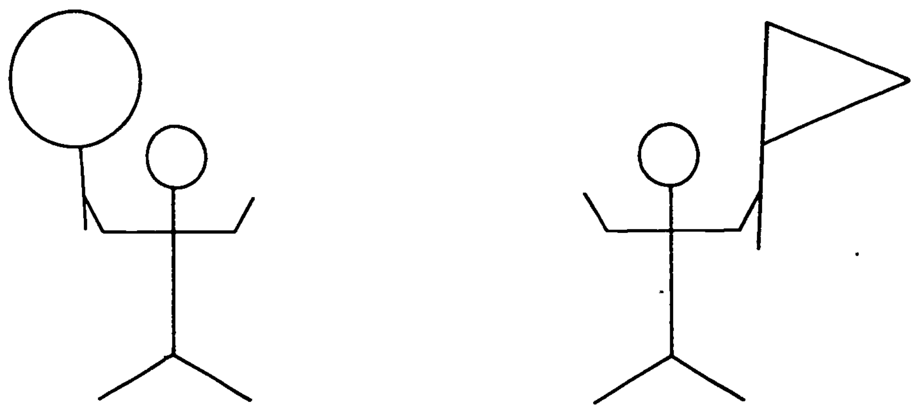
INCENTIVES IN EDUCATION
STUDENT QUESTIONNAIRE



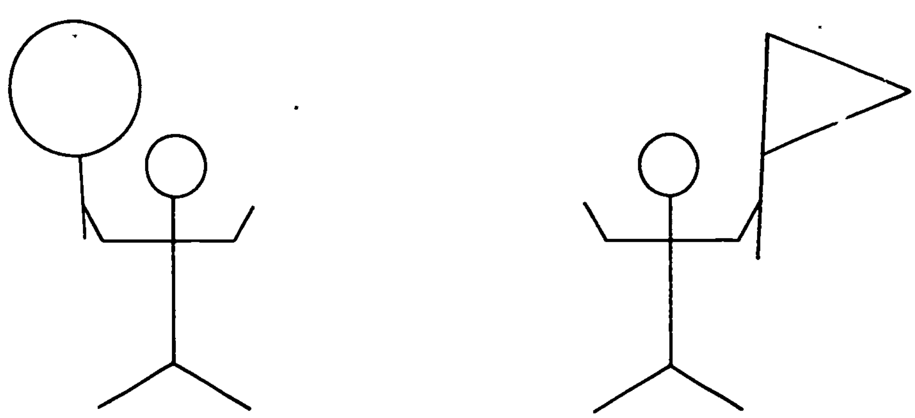
39. The flag-child wants to get training for a good job; the balloon-child does not care.



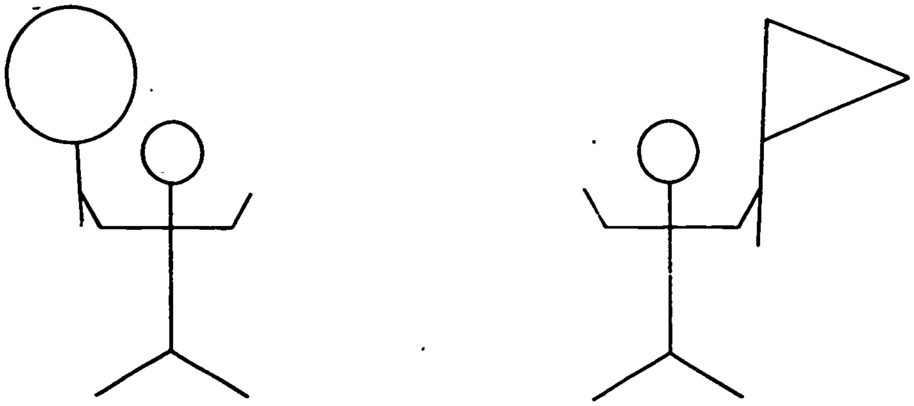
40. The balloon-child will get what he wants because of luck; the flag-child finishes his work.



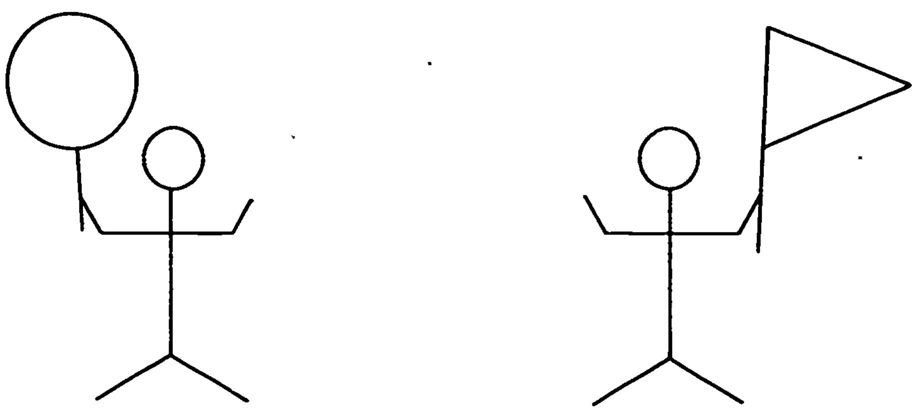
41. The balloon-child expects to be trained for a good job; the flag-child does not expect to be trained for a good job.



42. The balloon-child will get the job he wants. The flag-child will not get the job he wants.



43. Things are going to get worse for the balloon-child; things are going to get better for the flag-child.



OMB N° 51-57105S
Approval expires 7/31/72

NAME________________________ DISTRICT____________________ SCHOOL____________________

TEACHER QUESTIONNAIRE

This questionnaire contains two parts numbered Part I and Part II. The questions in Part I are to gather background information. Part II questions will be asked only on this first administration of the questionnaire. During the second administration later this year, you will not be required to complete Part II again.

DIRECTIONS: For each question, please place a check mark in front of the number of the response that is closest to your answer of place a number in the blank space provided. Answer all questions. Thank you.

PART I OF THE QUESTIONNAIRE ASKS ABOUT YOUR SCHOOL-RELATED ACTIVITIES.

1. **How much time per day (average) do you spend outside the normal school day on professional preparation (lesson planning, etc.)?**

   ________ hours

2. **What percentage of your time do you feel should go to:**

   top fourth of your students in arithmetic ________ %
   second fourth of your students in arithmetic ________ %
   third fourth of your students in arithmetic ________ %
   bottom fourth of your students in arithmetic ________ %

3. **What percentage of your time do you feel should go to:**

   top fourth of your students in reading ________ %
   second fourth of your students in reading ________ %
   third fourth of your students in reading ________ %
   bottom fourth of your students in reading ________ %

QUESTIONS 4 TO 28 ASK YOU ABOUT YOUR ATTITUDES TOWARD STUDENTS, TEACHING, AND THE SCHOOL. Please place a check mark in one of the boxes along the line to indicate how you feel about each question. There are no preferred answers to these questions. In the example below, the check mark shows that the person answering does not feel very strongly about television programs one way of the other.

EXAMPLE: I think television programs today are:

awful |____|____|__✓_|____|____| great

4. **My school:**

   could stand a lot of improvement |____|____|____|____|____| could serve as a model to others

5. Considering only their <u>motivation</u> to learn, my students are:
very poorly motivated |__|__|__|__|__| very highly motivated

6. Considering only their <u>academic ability</u>, my students are:
very limited |__|__|__|__|__| very capable

7. Considering only their <u>behavior</u>, my pupils behave:
very poorly |__|__|__|__|__| very well

8. I consider my work as a teacher:
not gratifying |__|__|__|__|__| gratifying

9. The offer of incentives to a teacher based on the achievement of his or her students is:
unlikely to increase achievement |__|__|__|__|__| likely to increase achievement

10. If I want to use new teaching methods or materials, my principal is likely to:
oppose me |__|__|__|__|__| assist me

11. The role most of my pupils' parents play in their children's education is:
of no significance |__|__|__|__|__| of great significance

12. My feeling about transferring to another school is:
very opposed |__|__|__|__|__| very favorable

13. My feeling about my status in this community (school attendance area) is:
dissatisfaction |__|__|__|__|__| satisfaction

14. Letting children move around in the classroom and talk to each other:
prevents learning |__|__|__|__|__| fosters learning

15. The offer of incentives to parents based on achievement of their child's class is:
not proper |__|__|__|__|__| proper

16. Please indicate your attitude toward your school:
strongly negative |__|__|__|__|__| strongly positive

17. Faculty meetings at my school are:
counterproductive |__|__|__|__|__| productive

18. Efforts of the faculty at my school to assist one another are:
counterproductive |__|__|__|__|__| productive

19. My principal's attitude towards new teaching methods and materials is:
negative |__|__|__|__|__| positive

20. Please indicate your attitude toward your pupils' parents:
strongly negative |__|__|__|__|__| strongly positive

21. The offer of incentives to a teacher based on the achievement of his or her students is:
not proper |__|__|__|__|__| proper

22. If a child is allowed to proceed at his own rate during a year's time, he is likely to:

learn less |___|___|___|___|___| learn more

23. My feelings about having faster pupils tutor slower pupils in my classroom are:

opposed |___|___|___|___|___| favorable

24. The offer of incentives to parents based on the achievement of their child's class is:

unlikely to increase achievement |___|___|___|___|___| likely to increase achievement

25. Strict discipline in the classroom is an important part of a child's education:

I strongly disagree |___|___|___|___|___| I strongly agree

26. Please indicate your attitude toward your pupils:

strongly negative |___|___|___|___|___| strongly positive

27. Having faster pupils tutor slower pupils is likely to be:

very ineffective |___|___|___|___|___| very effective

28. Individualized and self-paced instructions are likely to be:

very ineffective |___|___|___|___|___| very effective

PART II

DIRECTIONS: For each question, please place a check mark in front of the number of the response that is closest to your answer. Thank you.

1. What is your sex?
   ______ 1. Female
   ______ 2. Male

2. How old were you on your last birthday?
   ______ 1. under 26
   ______ 2. 26-35
   ______ 3. 36-45
   ______ 4. 46-55
   ______ 5. 56 or older

3. What will be your base-pay salary from this school system this year?
   ______ 1. $4000-4999
   ______ 2. $5000-5999
   ______ 3. $6000-6999
   ______ 4. $7000-7999
   ______ 5. $8000-8999
   ______ 6. $9000-9999
   ______ 7. $10,000-10,999
   ______ 8. $11,000-11,999
   ______ 9. $12,000-12,999
   ______10. $13,000 or more

4. What is your employment status in this school system?
   ______ 1. Tenured appointment
   ______ 2. Non-tenured appointment

5. What is the highest academic degree that you hold?
   ______ 1. None
   ______ 2. A degree based on less than 4 years of work
   ______ 3. A Bachelor's degree
   ______ 4. A Master's degree
   ______ 5. Professional or Specialist diploma
   ______ 6. A Doctorate

6. How many credits of graduate work have you had beyond your highest academic degree?
   ______ 1. None
   ______ 2. 1-10 credits
   ______ 3. 11-20 credits
   ______ 4. 21-30 credits
   ______ 5. 31+ credits

7. As of June, 1971, what was the total number of years that you have been a full time teacher?
   ______ years

8. How closely do you live to the school building in which you teach?
   ______ 1. Within the school building's normal attendance zone.
   ______ 2. Not in the building's attendance zone, but within one mile of the school.
   ______ 3. Inside the school district boundary.
   ______ 4. Outside of the school district.

9. How did you happen to be assigned to this particular school rather than some other school in this district?

_____ 1. I asked to work in this school.

_____ 2. I was placed in this school.

10. Did you request a transfer at any time from this school in school year 70-71?

_____ 1. This is my first year teaching in this building.

_____ 2. No

_____ 3. Yes

11. In the last year (70-71) approximately how many school days were you absent from this school? (Leave blank if you did not teach in this school last year).

_____days

OMB N° 51-571056
Approval expires 7/31/72

CHILD'S GRADE____________________ SCHOOL____________________ DISTRICT____________

PARENT QUESTIONNAIRE

PART I

INSTRUCTIONS: *This questionnaire is to be filled out by the person at home who is most responsible for the daily care and supervision of the child who brought this questionnaire home. Should more than one child in your family bring home a questionnaire please complete only one set. Whenever a question includes the phrase "your child," it means the child who brought home the questionnaire you are now completing. The questionnaire has two parts, numbered I and II. Part II will be asked only on the first questionnaire. Your cooperation is appreciated. Thank you.*

QUESTIONS 1-18 ASK ABOUT YOUR ATTITUDES TOWARD SCHOOL.

DIRECTIONS: *Please place a check mark in one of the boxes along the line to indicate how you feel about each question.*

There are no preferred answers to these questions.
In the example below, the check mark shows that the person answering does not feel very strongly about television programs one way or the other.

Example 1: I think television programs today are:

bad	no opinion	good
☐	☑	☐

In the example below, the check mark shows that the person answering does not like what is shown on television programs.

Example 2: I think television programs today are:

bad	no opinion	good
☑	☐	☐

In the below example, the check mark shows that the person answering does like what is shown on television programs.

Example 3: I think television programs today are:

bad	no opinion	good
☐	☐	☑

1. My child's school:

could stand improvement	no opinion	could serve as a model to others
☐	☐	☐

2. My child's teacher is:

one of the worst in the school	no opinion	one of the best in the school
☐	☐	☐

3. How do you feel about your child's school?

dissatisfied	no opinion	satisfied
☐	☐	☐

OE - 171-1

4. Offering teachers more money if their pupils learn more is:

not likely to work	no opinion	likely to work
☐	☐	☐

5. A good way to get my child to do well in school is to punish him when he does poorly.

I disagree	no opinion	I strongly agree
☐	☐	☐

6. When it comes to my child's education his teachers:

do not care	no opinion	do care a lot
☐	☐	☐

7. My feeling about discipline at my child's school is:

dissatisfied	no opinion	satisfied
☐	☐	☐

8. Encouraging parents with money if their child's class learns more is:

not proper	no opinion	proper
☐	☐	☐

9. To me, my child's ability to do arithmetic well:

does not matter	no opinion	is very important
☐	☐	☐

10. To me, my child's ability to read well:

does not matter	no opinion	is very important
☐	☐	☐

11. How do you feel about your child's teacher?

dissatisfied	no opinion	satisfied
☐	☐	☐

12. Offering teachers more money if their pupils learn more is:

not proper	no opinion	proper
☐	☐	☐

13. How is your child's teacher doing in teaching arithmetic to him?

poor job	no opinion	good job
☐	☐	☐

14. How is your child's teacher doing in teaching reading to him?

poor job	no opinion	good job
☐	☐	☐

15. A good way to get my child to do well in school is to praise him when he does well.

I disagree	no opinion	I agree
☐	☐	☐

16. Offering parents money if their child's class learns more is:

not likely to work	no opinion	likely to work
☐	☐	☐

17. How well is your child doing in arithmetic this year?

poorly	no opinion	well
☐	☐	☐

18. How well is your child doing in reading this year?

poorly	no opinion	well
☐	☐	☐

QUESTIONS 19 TO 33 ASK ABOUT YOUR INVOLVEMENT WITH YOUR CHILD'S EDUCATION.

DIRECTIONS: *Please circle the number of the response that is closest to your answer.*

19. How often does your child get help with arithmetic from someone who lives in your home (an adult or another child)?
    1. About every day
    2. About once a week
    3. About once a month
    4. A few times a year
    5. Never
    6. Don't know

20. How often does your child get help with arithmetic from an adult (other than the teacher) who does not live in your home?
    1. About every day
    2. About once a week
    3. About once a month
    4. A few times a year
    5. Never
    6. Don't know

21. How often does an adult in your home help other people's children with arithmetic?
    1. About every day
    2. About once a week
    3. About once a month
    4. A few times a year
    5. Never
    6. Don't know

22. How often does your child get help with arithmetic from another child?
    1. About every day
    2. About once a week
    3. About once a month
    4. A few times a year
    5. Never
    6. Don't know

23. How often does your child give help with arithmetic to another child?
    1. About every day
    2. About once a week
    3. About once a month
    4. A few times a year
    5. Never
    6. Don't know

24. How often do you encourage your child to do homework?
    1. About every day
    2. About once a week
    3. About once a month
    4. A few times a year
    5. Never
    6. Don't know

25. How often does your child get help with reading from someone who lives in your home (an adult or another child)?
    1. About every day
    2. About once a week
    3. About once a month
    4. A few times a year
    5. Never
    6. Don't know

26. How often does your child get help with reading from an adult (other than the teacher) who does not live in your home?
    1. About every day
    2. About once a week
    3. About once a month
    4. A few times a year
    5. Never
    6. Don't know

27. How often does an adult in your home help other people's children with reading?
    1. About every day
    2. About once a week
    3. About once a month
    4. A few times a year
    5. Never
    6. Don't know

28. How often does your child get help with reading from another child?
    1. About every day
    2. About once a week
    3. About once a month
    4. A few times a year
    5. Never
    6. Don't know

29. How often does your child give help with reading to another child?
    1. About every day
    2. About once a week
    3. About once a month
    4. A few times a year
    5. Never
    6. Don't know

30. When your child wants to do his schoolwork at home, can he have use of the following things (please place a check mark next to those things that he can use).
    1. A quiet room
    2. A desk or table
    3. A dictionary
    4. An encyclopedia set
    5. Blank paper
    6. Pencils, pens

31. About how many hours a week, if any, does your child watch television (please write in the number).
    ____________hours

32. About how many hours a week, if any, does your child watch the following children's television program (please write in the hours).
    Sesame Street ___________hours

33. About how many hours a week, if any, does your child watch the following children's television program (please write in the hours).
    The Electric Company __________hours

QUESTIONS 34 TO 38 ASK ABOUT YOUR COMMUNICATIONS WITH THE SCHOOL.

DIRECTIONS: Please place a number which approximates your activity for each question.

34. In the past two months, I have had about _____ meetings with my child's teacher. If no activity has taken place then mark a zero.

35. In the past two months, I have had about _____ telephone conversations with my child's teacher.

36. In the past two months, I have had about _____ written notes from my child's teacher.

37. In the past two months, I have been to _____ meetings of parents at the school.

38. In the past two months, I have been to _____ meetings with parents concerning school but not at the school.

PART II

QUESTIONS 1-7 ASK ABOUT YOU AND YOUR FAMILY.

DIRECTIONS: For each question, please circle the number of the response that is closest to your answer.

1. Excluding the child whose name appears on the first page how many children do you have at home?

   1. 1 to 2
   2. 3 to 4
   3. 5 to 6
   4. 7 to 8
   5. 9 or more

2. What is your relationship to your child whose name appears on the first page?
   1. Father
   2. Mother
   3. Foster parent
   4. Family relative
   5. Other

3. The approximate total family income for 1971 was (include all sources of income - work, gifts, welfare, social security and so on):
   1. 0-$2999
   2. $3000-4999
   3. $5000-6999
   4. $7000-8999
   5. $9000 or more

4. Please circle yes or no if you have the following things in your home.

1.	Yes	No	Television
2.	Yes	No	Radio
3.	Yes	No	Daily Newspaper
4.	Yes	No	Magazines
5.	Yes	No	Dictionary
6.	Yes	No	Encyclopedia
7.	Yes	No	Children's Books

5. How much schooling have you completed?
   1. Some grade school (less than 8 years)
   2. Completed grade school (8 years)
   3. Some high school (less than 4 years)
   4. Completed high school (4 years)
   5. Some vocational of business school after high school
   6. Completed vocational or business school after high school
   7. Some college
   8. Completed college
   9. Graduate or professional school

6. Select from the list below the category which best describes the job of the family bread winner.
   1. Unemployed
   2. Welfare recipient and unemployed
   3. Executives and proprietors of large concerns, major professionals (examples: doctor, lawyer, dentist, department store manager).
   4. Managers and proprietors of middle size businesses and service professionals (examples: policeman, fireman, branch manager of grocery chain).
   5. Administrative personnel of large concerns, owners of small businesses (examples: clothing shop owner, IBM programmer, florist).
   6. Clerical, retail sales workers, and technicians (examples: shipping clerk, dry cleaning shop attendant, x-ray technician).
   7. Skilled workers (examples: hair stylist, appliance repairman, electrician).
   8. Semi-skilled worker (examples: gas station attendant, delivery man, domestic worker).
   9. Unskilled workers (examples: cafeteria or laundry work, car washer, farm hand).
   10. If none of the above fits the description of his or her job, please describe the job in the space provided below.

____

____

7. What is your race/ethnic background?
   1. Black
   2. White, Spanish last name
   3. White
   4. Oriental
   5. American Indian
   6. Other

INCENTIVES IN EDUCATION

STUDENT INTERVIEW

Interviewer________________________

Date____________________

1. (a) Do you enjoy school? Tell me how you feel in the morning before coming to school--glad or not so glad.

P	FP	NO	FN	N

(b) What do you like or dislike about your school?

Record

2. (a) When you get to school do you like the work you do in class?

P	FP	NO	FN	N

(b) What is it you like or dislike about the work you do in school?

Record

3. (a) How do you feel about doing work in arithmetic, do you enjoy it or would you rather not do it?

P	FP	NO	FN	N

(b) What is it you like or dislike about arithmetic?

Record

4. (a) Do you enjoy work in reading, or would you rather not do it?

P	FP	NO	FN	N

(b) What is it you like or dislike about reading?

Record

5. (a) How many times do you visit the library on your own, out of class, once a week, twice a week, or more?

F	FF	OC	R	N
4	3	2	1	0

(b) What is it you like/dislike about your school library?

Record

6. (a) When you do your school work do you get any help from your friends, if so, how many times last week did you get help?

F	FF	OC	R	N
4	3	2	1	0

(b) Would you/do you like working with other students, or would you/do you prefer to work on your own?

Record

7. (a) Do you help your friends with their school work, if so, how many times last week did you give help?

F	FF	OC	R	N
4	3	2	1	0

(b) Would you/do you enjoy being able to help other students?

Record

8. (a) Would you like to be good at doing arithmetic problems or don't you care?

P	FP	NO	FN	N

(b) What is it you like/dislike about arithmetic problems?

Record

9. (a) Would you like to be a good reader?

P	FP	NO	FN	N

(b) What is it you like/dislike about reading?

Record

10. (a) When you finish school do you expect to get a good job?

P	FP	NO	FN	N

(b) What type of job would you like to have?

Record

11. (a) Does your teacher give you a hard time if your work is not good?

F	FF	OC	R	N
All time	Most time	Some time	Once while	

(b) What sort of things does your teacher say or do if your work is not good?

Record

12. (a) Do your parents, your mom or dad, give you a hard time if your work is not good?

F	FF	OC	R	N
All time	Most time	Some time	Once while	

(b) What sort of things do your mom and dad say or do if your work is not good?

Record

13. (a) Do you expect to be a good reader when you leave school?

P	FP	NO	PN	N

(b) What type of things do you want to be able to read, e.g. newspapers, magazines, books?

Record

14. (a) Do you expect to be good at arithmetic when you leave school?

P	FP	NO	PN	N

(b) What sort of things do you expect to be able to do with your arithmetic talent, e.g. say add up a shopping list, keep a check book, work in a store, teach, etc?

Record

15. (a) How do you get along with your teacher, well, or not so well, e.g. is she easy to talk to, can you ask her questions?

P	FP	NO	FN	N

(b) What is it you like/dislike about your teacher?

Record

INCENTIVES IN EDUCATION

TEACHER INTERVIEW

City____________________

Bldg.____________________

Tchr.____________________

Interviewer____________________

Interview No.____________________

Date____________________

1. (a) Do you think the payment of incentives to teachers is a good thing?

P	FP	NO	FN	N

(b) What specifically do you like/dislike about the idea of monetary incentives to teachers?

Record

2. (a) Do you think that the payment of incentives to parents is a good thing?

P	FP	NO	FN	N

(b) What specifically do you like/dislike about the idea of monetary incentives to parents?

Record

3. We would be pleased to receive your suggestions as to how this incentive project might have been improved. If you had been consulted in the early stages of the project what changes would you have made?

Record

4.(a) Have you made use of any techniques to individualize the curriculum you use in the classroom if so, how often in a normal school week would you use such techniques?

F	FF	OC	R	N
4	3	2	1	0

(b) Can you give me some examples of the sort of thing you have done?

Record

5.(a) How often in the last 2 months have you requested special materials (films, books, etc.), not normally supplied by the school district, to help you with your lessons?

F	FF	OC	R	N
4	3	2	1	0

(b) How successful were you/are you in getting such materials, and what were/are the major problems, if any?

Record

6.(a) Are you satisfied with the profession of teaching for you as an individual?

P	FP	NO	FN	N

(b) What are the specific likes/dislikes you have about your job as a teacher?

Record

7. (a) In the last 2 weeks have you ever considered transferring out of the school you are presently teaching in?

F	FF	OC	R	N
4	3	2	1	0

(b) What do you find particularly rewarding/unrewarding about teaching in this school?

Record

8. (a) In a normal school week do you use peer tutoring as a classroom teaching technique, if so, to what degree?

F	FF	OC	R	N
4	3	2	1	0

(b) Do you think that peer tutoring as a teaching technique could be usefully employed by the teacher?

Record

9. (a) Do you enjoy your students as children?

P	FP	NO	FN	N

(b) What do you particularly like or dislike about them?

Record

10. (a) What % of your students' parents have visited you in the last two months, and for what reasons?

F	FF	OC	R	N
16+	15-11	10-6	5-1	0

(b) Do you encourage parents in any way to visit with you and discuss their child's progress?

Record

11. (a) Do you feel you get on well with the parents of the children you teach?

F	FP	NO	FN	N

(b) What do you specifically like/dislike about the parents of the children you teach, and can you give any brief examples?

Record

12. (a) Many politicians and community members claim that the most important thing in the classroom should be discipline. Do you agree?

P	FP	NO	FN	N

(b) Do you find yourself faced with many discipline problems? Can you give some examples?

Record

13. (a) How do you feel about the ability of the parents of your pupils to help their children with their homework?

P	FP	NO	FN	N

(b) Have you any experience of parents helping their children with homework, and how successful was such aid?

Record

14. (a) How many hours, during a normal week, if it is necessary, do you spend on lesson preparation out of school and which subject areas take most time?

11+	10-7	6-4	3-1	0

(b) Do you find it necessary to spend much time on lesson preparation out of school time?

Record

15. In teaching reading to your students do you attempt to concentrate on any particular type of student, for e.g. the very bright, the very poor or those students in between? Can you elaborate on any strategies of this type that guide you in your teaching?

Record

16. In teaching arithmetic to your students do you attempt to concentrate on any particular type of student, for e.g. the very bright, the very poor, of those students in between? Can you elaborate on any strategies of this type that guide you in your teaching?

Record

17. How much use does your class make of the school and public libraries? Do you find it necessary to offer your students encouragement in order that they use the library; if so, what form does it take?

Record

18. (a) What are your feelings towards the incentive project being tried in your school district?

P FP NO FN N

(b) What in particular do you like/dislike about the project?

Record

19. Do you have any suggestions for any other forms of incentives that might be used to encourage academic achievement gains in students?

Record

INCENTIVES IN EDUCATION

PARENT INTERVIEW

Parent Name________________________

Interviewer________________________

Relation to child______________ Date______

1. (a) Are you happy with the school your child is attending this year?

   (b) What is it you like/dislike, are happy/unhappy with about the school?

   P FP NO FN N

2. (a) Do you attend school activities, e.g. PTA meetings, or other school activities?

* (b) How many times have you visited the school in the last last two months?

   F FF OC R N
   4 3 2 1 0

3. (a) How often do you encourage your child to do well at his school work?

* (b) How many times during the past week have you offered a form of encouragement?

   F FF OC R N
   4 3 2 1 0

4. (a) Do you help your child in doing his/her homework?

* (b) How many times <u>last week</u> did you help your child in his/her homework?

   F FF OC R N
   4 3 2 1 0

5. (a) Do you think that groups of parents working together can improve their children's school grades?

   (b) Have you had any experiences that make you feel that way?

   P FP NO FN N

6. (a) Have you organized any extra-help sessions for reading/arithmetic work for your children, with other parents in the neighborhood?

* (b) How often in the last month have such sessions been held?

F	FF	OC	R	N
4	3	2	1	0

(c) Who has taken the lead, to your knowledge, in organizing parent groups? Can you give me any names of parents you know?

Record

7. (a) Are you satisfied with how your child is doing in arithmetic?

(b) What do you feel is good/bad about what is being taught?

P	FP	NO	FN	N

8. (a) Are you satisfied with how your child is doing in reading?

(b) What do you feel is good/bad about what is being taught?

P	FP	NO	FN	N

9. (a) Do you get extra reading and/or arithmetic learning materials for your child?

* (b) How many materials have you obtained in the last 2 months?

F	FF	OC	R	N
3	2	1	Yes but not in the last 2 months	0

10. (a) Do you encourage your child to use the public library?

* (b) How often have you encouraged your child in this way in the last two weeks?

F FF OC R N

4 3 2 1 0

11. (a) How often does your child use the public/school library?

* (b) How often has he brought books home in the last two weeks?

F FF OC R N

4 3 2 1 0

12. (a) Do you think parents need to be familiar with their child's teacher?

(b) Do you feel you know your child's teacher as well as you would like?

P FP NO FN N

13. (a) Have you visited with or met your child's teacher?

* (b) How many times in the last two months have you met or talked with your child's teacher?

F FF OC R N

4 3 2 1 0

14. (a) Do your child's friends work with your child on their reading/arithmetic homework?

* (b) How often has that occurred in the last two weeks?

F FF OC R N

4 3 2 1 0

15. (a) Do any adults outside of your home help your child with his/her homework?

* (b) How often in the last two months has such help been given?

F FF OC R N

4 3 2 1 0

16. (a) Do you approve of the way the school is teaching your child to behave?

(b) What particularly do you like/dislike?

P	FP	NO	FN	N

17. (a) Do you approve the idea of payment of money being paid to teachers, e.g. a bonus, if your child learns more?

(b) What do you particularly like/dislike?

P	FP	NO	FN	N

18. (a) Do you like the idea of payments of money being paid to parents if their child learns more due to parental help?

(b) What particularly do you like/dislike about the idea?

P	FP	NO	FN	N

19. (a) Have you discussed or been to meetings with other parents about the payments of money project?

* (b) How many times have you had such discussions/ meetings in the last two weeks?

F	FF	OC	R	N
4	3	2	1	0

20. (a) Have you discussed or attended meetings about the payments of money project with your child's teacher or other school official?

* (b) How many times have you had such discussions/ meetings in the last two weeks?

F	FF	OC	R	N
4	3	2	1	0

21. If you could have had a say in setting this project up, how would you have done it differently?

Record

22. Do you have any suggestions for other forms of incentives that might be used to stimulate activity and learning, in the classroom?

Record

APPENDIX B

TABLES OF ACHIEVEMENT DATA RESULTS
GRADE BY GRADE

TABLE B-1

READ STAN GAIN

CINCINNATI								
GRADE	MEAN 1	SIGMA 1	NUM 1	MEAN 2	SIGMA 2	NUM 2	t-VALUE	R-SQUARED
1	8.3864	4.1271	44.0	9.9219	4.2810	64.0	-1.8583	.0316
2	8.3333	4.7907	48.0	10.2778	5.3214	54.0	-1.9299	.0359
3	5.6222	6.0988	45.0	5.4262	3.8878	61.0	+0.2017	.0004
4	5.4021	5.1774	97.0	4.4063	4.3597	64.0	+1.2698	.0100
5	5.4231	4.2197	78.0	3.0339	6.5862	59.0	+2.5806	.0470
6	3.7835	5.5775	97.0	5.5167	6.6243	60.0	-1.7595	.0196

JACKSONVILLE								
GRADE	MEAN 1	SIGMA 1	NUM 1	MEAN 2	SIGMA 2	NUM 2	t-VALUE	R-SQUARED
1	9.3750	4.0739	112.0	10.3506	4.3279	77.0	-1.5770	.0131
2	6.5865	5.2644	133.0	10.0886	6.1231	79.0	-4.4037	.0845
3	7.7231	5.4956	130.0	5.7742	5.0025	93.0	+2.7096	.0322
4	6.2710	6.6851	107.0	4.9524	4.8230	126.0	+1.7436	.0130
5	4.4420	5.9878	138.0	4.4797	5.0702	148.0	-0.0576	.0000
6	7.4492	7.2861	118.0	6.1471	5.5263	170.0	+1.7233	.0103

TABLE B-1 (Cont'd.)

				SA.. ANTONIO				
GRADE	MEAN 1	SIGMA1	NUM 1	MEAN 2	SIGMA 2	NUM 2	t-VALUE	R-SQUARED
1	10.5065	5.4619	77.0	14.7273	6.9829	44.0	-3.6881	.1026
2	10.0625	5.9830	16.0	8.7045	4.8780	44.0	+0.8969	.0137
3	7.3944	5.4941	71.0	6.3607	4.0948	61.0	+1.2088	.0111
4	7.3571	7.7938	56.0	6.1154	6.4585	52.0	+0.8977	.0075
5	5.3542	9.2977	48.0	5.4222	7.2252	45.0	-0.0392	.0000
6	8.0615	7.6053	65.0	5.2857	5.4740	42.0	+2.0462	.0383

				OAKLAND				
GRADE	MEAN 1	SIGMA1	NUM 1	MEAN 2	SIGMA 2	NUM 2	t-VALUE	R-SQUARED
1	10.2381	5.8440	63.0	9.0175	4.5845	57.0	+1.2636	.0134
2	10.7377	6.1479	61.0	5.6538	6.3119	78.0	+4.7662	.1422
3	7.1087	4.6678	46.0	4.2870	5.3233	108.0	+3.1192	.0602
4	4.6966	5.4050	89.0	5.8116	5.9663	69.0	-1.2289	.0096
5	3.6863	4.6454	51.0	4.1250	6.1979	80.0	-0.4336	.0015
6	6.7843	6.1198	51.0	3.2105	5.0632	95.0	+3.7752	.0901

TABLE B-2

MATH STAN GAIN

	CINCINNATI							
GRADE	MEAN 1	SIGMA 1	NUM 1	MEAN 2	SIGMA 2	NUM 2	t-VALUE	R-SQUARED
1	15.8636	7.6206	44.0	15.1538	8.3801	65.0	+0.4498	.0019
2	15.8980	8.1400	49.0	15.0800	7.0994	50.0	+0.5332	.0029
3	5.5111	9.0745	45.0	9.1803	6.2544	61.0	-2.4644	.0552
4	8.5567	6.6536	97.0	9.3810	6.6099	63.0	-0.7676	.0037
5	6.9359	5.9157	78.0	4.8644	6.3990	59.0	+1.9592	.0276
6	3.8247	4.9456	97.0	4.1500	5.8363	60.0	-0.3735	.0009

	JACKSONVILLE							
GRADE	MEAN 1	SIGMA 1	NUM 1	MEAN 2	SIGMA 2	NUM 2	t-VALUE	R-SQUARED
1	15.8319	8.1579	113.0	16.4416	6.8739	77.0	-0.5383	.0015
2	11.4885	8.4747	131.0	13.3590	7.5918	78.0	-1.6032	.0123
3	9.8682	7.6110	129.0	7.8901	8.0090	91.0	+1.8578	.0156
4	9.3619	6.1082	105.0	8.1453	6.5182	117.0	+1.4303	.0092
5	6.5374	5.3098	138.0	6.5374	5.7616	147.0	-0.4206	.0006
6	7.2288	5.7041	118.0	4.4024	4.2836	169.0	+4.7920	.0746

TABLE B-2 (Cont'd.)

OAKLAND								
GRADE	MEAN 1	SIGMA 1	NUM 1	MEAN 2	SIGMA 2	NUM 2	t-VALUE	R-SQUARED
1	15.9048	7.8734	63.0	11.9298	8.0264	57.0	+2.7364	.0597
2	14.4167	6.4500	60.0	11.4359	8.9000	78.0	+2.1889	.0340
3	10.2889	6.3160	45.0	6.6944	6.9153	108.0	+3.0029	.0564
4	7.0341	6.9024	88.0	6.2857	7.3171	70.0	+0.6592	.0028
5	4.8824	5.4430	51.0	3.4000	6.4053	80.0	+1.3673	.0143
6	7.0600	5.4037	50.0	2.9053	5.1075	95.0	+4.5634	.1271

SAN ANTONIO								
GRADE	MEAN 1	SIGMA 1	NUM 1	MEAN 2	SIGMA 2	NUM 2	t-VALUE	R-SQUARED
1	22.3636	11.7920	77.0	17.0455	7.7127	44.0	+2.6795	.0569
2	11.8125	6.0797	16.0	12.3182	7.6759	44.0	-0.2374	.0010
3	9.1831	6.5950	71.0	6.4262	8.0093	61.0	+2.1680	.0349
4	13.4464	6.7012	56.0	5.6538	7.0984	52.0	+5.8684	.2452
5	8.0816	7.1439	49.0	6.4773	4.9860	44.0	+1.2324	.0167
6	8.2615	7.4084	65.0	6.5476	4.4018	42.0	+1.3517	.0171

TABLE B-3

READ STAN PRE AND POST

CINCINNATI PRE

	CINCINNATI							
GRADE	MEAN 1	SIGMA 1	NUM 1	MEAN 2	SIGMA 2	NUM 2	t-VALUE	R-SQUARED
1	26.2727	4.5257	44.0	24.6406	3.5740	64.0	+2.0900	.0396
2	34.1250	6.3300	48.0	35.5370	8.3750	54.0	-0.9511	.0090
3	47.2667	10.2523	45.0	47.9836	7.3813	61.0	-0.4188	.0017
4	50.4639	9.2512	97.0	53.8281	9.0423	64.0	-2.2784	.0316
5	55.7564	11.7939	78.0	60.2373	8.4228	59.0	-2.4782	.0435
6	63.2165	12.5874	97.0	60.9833	12.0613	60.0	+1.0974	.0077

CINCINNATI POST

	CINCINNATI							
GRADE	MEAN 1	SIGMA 1	NUM 1	MEAN 2	SIGMA 2	NUM 2	t-VALUE	R-SQUARED
1	34.6591	6.1830	44.0	34.5625	5.4274	64.0	+0.0858	.0001
2	42.4583	9.0294	48.0	45.8148	6.9908	54.0	-2.1114	.0427
3	52.8889	9.8170	45.0	53.4098	7.4931	61.0	-0.3099	.0009
4	55.8660	9.8156	97.0	58.2344	9.4948	64.0	-1.5178	.0143
5	61.1795	12.7266	78.0	63.2712	10.9463	59.0	-1.0107	.0075
6	67.0000	13.8346	97.0	66.5000	11.3055	60.0	+0.2354	.0004

TABLE B-3 (Cont'd.)

JACKSONVILLE PRE

JACKSONVILLE								
GRADE	MEAN 1	SIGMA 1	NUM 1	MEAN 2	SIGMA 2	NUM 2	t-VALUE	R-SQUARED
1	22.3929	4.8255	112.0	21.8701	5.1564	77.0	+0.7115	.0027
2	28.7669	6.1483	133.0	32.0127	6.3841	79.0	-3.6637	.0601
3	35.3615	7.7660	130.0	41.8065	9.6169	93.0	-5.5276	.1215
4	42.8879	12.1207	107.0	48.3651	9.7444	126.0	-3.8226	.0595
5	51.6014	11.4398	138.0	52.7432	10.0895	148.0	-0.8966	.0028
6	54.6356	14.4835	118.0	59.3647	12.0829	170.0	-3.0087	.0307

JACKSONVILLE POST

JACKSONVILLE								
GRADE	MEAN 1	SIGMA 1	NUM 1	MEAN 2	SIGMA 2	NUM 2	t-VALUE	R-SQUARED
1	31.7679	4.3949	112.0	32.2208	5.8075	77.0	-0.6098	.0020
2	35.3534	8.3713	133.0	42.1013	9.9854	79.0	-5.2756	.1170
3	43.0846	8.9443	130.0	47.5806	10.7383	93.0	-3.4018	.0498
4	49.1589	12.7102	107.0	53.3175	10.5061	126.0	-2.7341	.0313
5	56.0435	13.4147	138.0	57.2230	11.6203	148.0	-0.7962	.0022
6	62.0847	15.5517	118.0	65.5118	14.0953	170.0	-1.9445	.0130

TABLE B-3 (Cont'd.)

OAKLAND PRE

	OAKLAND							
GRADE	MEAN 1	SIGMA1	NUM 1	MEAN 2	SIGMA 2	NUM 2	t-VALUE	R-SQUARED
1	26.0000	4.8957	63.0	27.5088	4.1193	57.0	-1.816	.0272
2	36.4754	8.6170	61.0	36.2051	8.4306	78.0	+0.1858	.0003
3	49.4783	8.3898	46.0	49.9537	9.4150	108.0	-0.2960	.0006
4	51.3034	12.4082	89.0	56.1739	9.4011	69.0	-2.7118	.0450
5	58.1765	10.7419	51.0	60.9750	12.1749	80.0	-1.3417	.0138
6	68.7451	15.2156	51.0	68.5789	10.4256	95.0	+0.0778	.0000

OAKLAND POST

	OAKLAND							
GRADE	MEAN 1	SIGMA1	NUM 1	MEAN 2	SIGMA 2	NUM 2	t-VALUE	R-SQUARED
1	36.2381	5.7043	63.0	36.5263	4.6679	57.0	-0.3010	.0008
2	47.2131	8.1017	61.0	41.8590	7.1926	78.0	+4.1195	.1102
3	56.5870	8.1937	46.0	54.2407	8.9588	108.0	+1.5248	.0151
4	56.0000	12.7618	89.0	61.9855	12.8252	69.0	-2.9177	.0517
5	61.8627	12.1557	51.0	65.1000	10.9077	80.0	-1.5837	.0191
6	75.5294	13.9561	51.0	71.7895	11.1354	95.0	+1.7675	.0212

TABLE B-3 (Cont'd.)

SAN ANTONIO PRE

	SAN ANTO.IO							
GRADE	MEAN 1	SIGMA1	NUM 1	MEAN 2	SIGMA 2	NUM 2	t-VALUE	R-SQUARED
1	24.8701	4.1082	77.0	22.5000	4.6929	44.0	+2.8974	.0659
2	27.6875	5.1990	16.0	32.4773	5.1466	44.0	-3.1795	.1484
3	43.2254	10.9299	71.0	45.2623	9.7637	61.0	-1.1210	.0096
4	49.8036	9.5678	56.0	52.7308	11.3865	52.0	-1.4500	.0195
5	55.3542	13.2705	48.0	62.8000	B.7024	45.0	-2.6618	.0722
6	59.8769	12.3876	65.0	70.6190	9.8227	42.0	-4.7369	.1761

SAN ANTONIO POST

	SAN ANTONIO							
GRADE	MEAN 1	SIGMA1	NUM 1	MEAN 2	SIGMA 2	NUM 2	t-VALUE	R-SQUARED
1	35.3766	5.4048	77.0	37.2273	5.7258	44.0	-1.7731	.0257
2	37.7500	6.6383	16.0	41.1818	7.5491	44.0	-1.6049	.0425
3	50.6197	12.0823	71.0	51.6230	9.4836	61.0	-0.5243	.0021
4	57.1607	8.9112	56.0	58.8462	12.9923	52.0	-0.7910	.0059
5	60.7083	13.7469	48.0	68.2222	14.3272	45.0	-2.5809	.0682
6	67.9385	13.6426	65.0	75.9048	11.2549	42.0	-3.1526	.0865

TABLE B-4

MATH STAN PRE AND POST

CINCINNATI PRE

	CINCINNATI							
GRADE	MEAN 1	SIGMA 1	NUM 1	MEAN 2	SIGMA 2	NUM 2	t-VALUE	R-SQUARED
1	23.1591	6.3207	44.0	22.1077	5.2154	65.0	+0.9473	.0083
2	33.8367	5.3942	49.0	36.1400	9.7248	50.0	-1.4532	.0213
3	46.5556	13.5740	45.0	51.1311	10.4920	61.0	-1.9577	.0355
4	55.6495	12.1793	97.0	57.7143	10.2979	63.0	-1.1118	.0078
5	66.1667	11.3531	78.0	68.4746	8.0502	59.0	-1.3286	.0129
6	72.2784	11.2997	97.0	76.8500	10.3003	60.0	-2.5466	.0402

CINCINNATI POST

	CINCINNATI							
GRADE	MEAN 1	SIGMA 1	NUM 1	MEAN 2	SIGMA 2	NUM 2	t-VALUE	R-SQUARED
1	39.0227	11.6349	44.0	37.2615	11.6113	65.0	+0.7763	.0056
2	49.7347	10.4458	49.0	51.2200	11.7635	50.0	-0.6638	.0045
3	52.0667	12.2147	45.0	60.3115	11.2480	61.0	-3.5962	.1106
4	64.2062	13.3432	97.0	67.0952	13.1722	63.0	-1.3448	.0113
5	73.1026	11.8086	78.0	73.3390	8.9532	59.0	-0.1284	.0001
6	76.1031	11.9412	97.0	81.0000	12.5333	60.0	-2.4499	.0373

TABLE B-4 (Cont'd.)

JACKSONVILLE PRE

JACKSONVILLE								
GRADE	MEAN 1	SIGMA 1	NUM 1	MEAN 2	SIGMA 2	NUM 2	t-VALUE	R-SQUARED
1	18.2124	5.4714	113.0	17.0000	6.3702	77.0	+1.4021	.0103
2	29.3511	7.8230	131.0	29.8846	7.8773	78.0	-0.4756	.0011
3	36.7442	9.1244	129.0	43.4615	9.9580	91.0	-5.1774	.1095
4	48.5905	12.0517	105.0	51.8462	11.4339	117.0	-2.0647	.0190
5	58.7681	12.4633	138.0	61.7959	12.7553	147.0	-2.0250	.0143
6	69.1271	14.3268	118.0	71.3787	12.5389	169.0	-1.4110	.0069

JACKSONVILLE POST

JACKSONVILLE								
GRADE	MEAN 1	SIGMA 1	NUM 1	MEAN 2	SIGMA 2	NUM 2	t-VALUE	R-SQUARED
1	34.0442	8.9297	113.0	33.4416	10.1430	77.0	+0.4321	.0010
2	40.8397	12.2593	131.0	43.2436	11.2852	78.0	-1.4117	.0095
3	46.6124	11.2592	129.0	51.3516	13.1727	91.0	-2.8644	.0363
4	57.9524	13.9404	[illegible]	59.9915	12.1673	117.0	-1.1636	.0061
5	65.0290	13.4858	138.0	68.3333	13.0905	147.0	-2.0987	.0153
6	76.3559	15.6977	118.0	75.7811	13.4790	169.0	+0.3321	.0004

TABLE B-4 (Cont'd.)

OAKLAND PRE

	OAKLAND							
GRADE	MEAN 1	SIGMA1	NUM 1	MEAN 2	SIGMA 2	NUM 2	t-VALUE	R-SQUARED
1	21.7937	4.9192	63.0	25.2632	4.6657	57.0	-3.9536	.1170
2	36.6667	9.1719	60.0	35.8718	9.1982	78.0	+0.5039	.0019
3	51.8889	8.5579	45.0	50.8981	11.3524	108.0	+0.5261	.0018
4	54.9318	14.4991	88.0	60.2000	12.3060	70.0	-2.4235	.0363
5	62.1961	12.1277	51.0	70.6250	13.5173	80.0	-3.6195	.0922
6	78.7800	13.3314	50.0	77.5895	12.0711	95.0	+0.5444	.0021

OAKLAND POST

	OAKLAND							
GRADE	MEAN 1	SIGMA1	NUM 1	MEAN 2	SIGMA 2	NUM 2	t-VALUE	R-SQUARED
1	37.6984	9.0708	63.0	37.1930	9.2185	57.0	+0.3025	.0008
2	51.0833	11.1085	60.0	47.3077	10.3426	78.0	+2.0584	.0302
3	62.1778	9.4515	45.0	57.5926	11.9203	108.0	+2.2957	.0337
4	61.9659	12.7950	88.0	66.4857	13.4862	70.0	-2.1535	.0289
5	67.0784	13.2843	51.0	74.0250	12.3555	80.0	-3.0469	.0671
6	85.8400	12.1594	50.0	80.4947	11.0134	95.0	+2.6792	.0478

TABLE B-4 (Cont'd.)

SAN ANTONIO PRE

	SAN ANTONIO							
GRADE	MEAN 1	SIGMA1	NUM 1	MEAN 2	SIGMA 2	NUM 2	t-VALUE	R-SQUARED
1	22.8831	6.7609	77.0	21.7955	6.1553	44.0	+0.8789	.0064
2	31.3125	4.8404	16.0	35.4773	6.4861	44.0	-2.3375	.0861
3	47.8310	12.5732	71.0	50.5410	11.0839	61.0	-1.3035	.0129
4	50.7857	11.7206	56.0	58.3269	12.3220	52.0	-3.2595	.0911
5	65.6939	11.8677	49.0	72.7727	11.7748	44.0	-2.8826	.0837
6	71.3538	14.3390	65.0	78.5476	9.8233	42.0	-2.8461	.0716

SAN ANTONIO POST

	SAN ANTONIO							
GRADE	MEAN 1	SIGMA1	NUM 1	MEAN 2	SIGMA 2	NUM 2	t-VALUE	R-SQUARED
1	45.2468	11.3244	77.0	38.8409	7.4707	44.0	+3.3551	.0864
2	43.1250	6.8593	16.0	47.7955	9.7708	44.0	-1.7566	.0505
3	57.0141	14.0667	71.0	56.9672	11.9345	61.0	+0.0205	.0000
4	64.2321	9.2402	56.0	63.9808	13.4725	52.0	+0.1138	.0001
5	73.7755	11.7052	49.0	79.2500	13.3454	44.0	-2.1075	.0465
6	79.6154	13.8932	65.0	85.0952	8.5762	42.0	-2.2879	.0475